

Best Local Similarity 100.0%; Pred. No. 1.7e-90;
Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSGLGRSRGGRSRVDQERFPQGLWTGVAMRSCPEEQYWDPLLTGTCMCKTICNHQSQR 60
Db 1 MSGLGRSRGGRSRVDQERFPQGLWTGVAMRSCPEEQYWDPLLTGTCMCKTICNHQSQR 60

Qy 61 TCAAFCSRSLSCRKEQKGFYDHLRDCISCASTCGQHPKQCAFCENKLRSPVNLPPELRR 120
Db 61 TCAAFCSRSLSCRKEQKGFYDHLRDCISCASTCGQHPKQCAFCENKLRSPVNLPPELRR 120

Qy 121 QRSGEVNNSDNSGRYQGLEHRGSEASPALPGLKLSADQVALVYST 166
Db 121 QRSGEVNNSDNSGRYQGLEHRGSEASPALPGLKLSADQVALVYST 166

RESULT 2
US-09-290-333-6
; Sequence 6, Application US/09290333
; Patent No. 6316222
; GENERAL INFORMATION:
; APPLICANT: Bram, Richard J.
; von Bulow, Gotz
; TITLE OF INVENTION: A LYMPHOCYTE SURFACE RECEPTOR THAT BINDS
; CAML, NUCLEIC ACIDS ENCODING THE SAME AND METHODS OF USE
; THEREOF
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th
; Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/290,333
; FILING DATE: 12-Apr-1999
; CLASSIFICATION: <unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 1340-1-007 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 166 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHEICAL: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:

US-09-290-333-6

Query Match 100.0%; Score 909; DB 4; Length 166;
Best Local Similarity 100.0%; Pred. No. 1.7e-90;
Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSGLGRSRGGRSRVDQERFPQGLWTGVAMRSCPEEQYWDPLLTGTCMCKTICNHQSQR 60
Db 1 MSGLGRSRGGRSRVDQERFPQGLWTGVAMRSCPEEQYWDPLLTGTCMCKTICNHQSQR 60

Qy 61 TCAAFCSRSLSCRKEQKGFYDHLRDCISCASTCGQHPKQCAFCENKLRSPVNLPPELRR 120
Db 61 TCAAFCSRSLSCRKEQKGFYDHLRDCISCASTCGQHPKQCAFCENKLRSPVNLPPELRR 120

Qy 121 QRSGEVNNSDNSGRYQGLEHRGSEASPALPGLKLSADQVALVYST 166
Db 121 QRSGEVNNSDNSGRYQGLEHRGSEASPALPGLKLSADQVALVYST 166

RESULT 3
US-08-810-572A-2
; Sequence 2, Application US/08810572A
; Patent No. 5969102
; GENERAL INFORMATION:
; APPLICANT: Bram, Richard J.
; von Bulow, Gotz
; TITLE OF INVENTION: A LYMPHOCYTE SURFACE RECEPTOR THAT BINDS
; CAML, NUCLEIC ACIDS ENCODING THE SAME AND METHODS OF USE
; THEREOF
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th
; Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/810,572A
; FILING DATE: 28-FEB-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 1340-1-007
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 293 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; US-08-810-572A-2

Query Match 100.0%; Score 909; DB 2; Length 293;
Best Local Similarity 100.0%; Pred. No. 3.4e-50;
Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSGLGRSRGGRSRVDQERFPQGLWTGVAMRSCPEEQYWDPLLTGTCMCKTICNHQSQR 60
Db 1 MSGLGRSRGGRSRVDQERFPQGLWTGVAMRSCPEEQYWDPLLTGTCMCKTICNHQSQR 60

Qy 61 TCAAFCSRSLSCRKEQKGFYDHLRDCISCASTCGQHPKQCAFCENKLRSPVNLPPELRR 120
Db 61 TCAAFCSRSLSCRKEQKGFYDHLRDCISCASTCGQHPKQCAFCENKLRSPVNLPPELRR 120

Qy 121 QRSGEVNNSDNSGRYQGLEHRGSEASPALPGLKLSADQVALVYST 166
Db 121 QRSGEVNNSDNSGRYQGLEHRGSEASPALPGLKLSADQVALVYST 166

Db 121 QRSGEVENNSDNGRYQGLEHRGSEASPALPGLKLSADQVALVYST 166

RESULT 4

US-09-290-333-2

; Sequence 2, Application US/09290333

; Patent No. 6316222

; GENERAL INFORMATION:

; APPLICANT: Bram, Richard J.

; von Bulow, Gotz

; TITLE OF INVENTION: A LYMPHOCYTE SURFACE RECEPTOR THAT BINDS

; CAML, NUCLEIC ACIDS ENCODING THE SAME AND METHODS OF USE

; THEREOF

; NUMBER OF SEQUENCES: 11

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: David A. Jackson, Esq.

; STREET: 411 Hackensack Ave, Continental Plaza, 4th

; Floor

; CITY: Hackensack

; STATE: New Jersey

; COUNTRY: USA

; ZIP: 07601

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/290.333

; FILING DATE: 12-Apr-1999

; CLASSIFICATION: <Unknown>

; ATTORNEY/AGENT INFORMATION:

; NAME: Jackson Esq., David A.

; REGISTRATION NUMBER: 26,742

; REFERENCE/DOCKET NUMBER: 1340-1-007 PCT

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 201-487-5800

; TELEFAX: 201-343-1684

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 293 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; HYPOTHETICAL: NO

; FRAGMENT TYPE: N-terminal

; ORIGINAL SOURCE:

; ORGANISM: Homo sapiens

; SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-290-333-2

Query Match 100.0%; Score 909; DB 4; Length 293;

Best Local Similarity 100.0%; Pred. No. 3.4e-90;

Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSGLGRSRGGRSRVDQERPPQGLWTGVAMRSCPEEQYWDPLLGTCMSCKTICNHQSQR 60

|||||

Db 1 MSGLGRSRGGRSRVDQERPPQGLWTGVAMRSCPEEQYWDPLLGTCMSCKTICNHQSQR 60

|||||

Qy 61 TCAAFCSLSCKRQGGKPYDHLRDCISCAICGQHPKQCAIFCENKLRSPVNPPELRR 120

|||||

Db 61 TCAAFCSLSCKRQGGKPYDHLRDCISCAICGQHPKQCAIFCENKLRSPVNPPELRR 120

|||||

Qy 121 QRSGEVENNSDNGRYQGLEHRGSEASPALPGLKLSADQVALVYST 166

|||||

Db 121 QRSGEVENNSDNGRYQGLEHRGSEASPALPGLKLSADQVALVYST 166

|||||

RESULT 5

US-08-525-940-23

; Sequence 23, Application US/08525940

; Patent No. 5866351

; GENERAL INFORMATION:

; APPLICANT: Franzusoff, Alex

; APPLICANT: Miranda, Luis R.

; APPLICANT: Wolf, Joseph R.

; TITLE OF INVENTION: CD4+ T-LYMPHOCYTE PROTEASES AND GENES

; TITLE OF INVENTION: ENCODING SAID PROTEASES

; NUMBER OF SEQUENCES: 25

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Sheridan Ross & McIntosh

; STREET: 1700 Lincoln Street, Suite 3500

; CITY: Denver

; STATE: Colorado

; COUNTRY: U.S.A.

; ZIP: 80203

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/525,940

; FILING DATE:

; CLASSIFICATION: 514

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/368,852

; FILING DATE: 01-JAN-1995

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/088,322

; FILING DATE: 07-JUL-1993

; ATTORNEY/AGENT INFORMATION:

; NAME: Connell, Gary J.

; REGISTRATION NUMBER: 32,020

; REFERENCE/DOCKET NUMBER: 2848-11-C1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (303) 863-9700

; TELEFAX: (303) 863-0223

; INFORMATION FOR SEQ ID NO: 23:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 799 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-525-940-23

Query Match 9.1%; Score 82.5; DB 2; Length 799;

Best Local Similarity 26.0%; Pred. No. 1.3;

Matches 27; Conservative 10; Mismatches 34; Indels 33; Gaps 7;

Qy 27 TGVMRSCPEEQYWDPLLGTCMSCKTICNHQSORTCAAF-----CR-----SLSC 71

Db 603 TNSCVTHCPDGSYQDTKKNLCRKCSENC-----KTCTEFHNCTECRDGLSLQGRCSYSC 657

|||||

Qy 72 RKEQKDYDHLRDCISCAIC-----GQHPKQC-----AYFCEN 106

|||||

Db 658 --EDGRYFNG--QDCQCPCHRCATCAGAGAGCINCTEGYFME 697

|||||

RESULT 6

US-08-976-838-23

; Sequence 23, Application US/08976838

; Patent No. 5981259

; GENERAL INFORMATION:

; APPLICANT: Franzusoff, Alex

; TITLE OF INVENTION: CD4+ T-LYMPHOCYTE PROTEASE NUCLEIC ACID

; TITLE OF INVENTION: MOLECULES

; NUMBER OF SEQUENCES: 31

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Sheridan Ross P.C.

; STREET: 1700 Lincoln St., Suite 3500

; CITY: Denver

; STATE: Colorado

```

; ATTORNEY/AGENT INFORMATION:
; NAME: Connell, Gary J.
; REGISTRATION NUMBER: 32,020
; REFERENCE/DOCKET NUMBER: 2848-11-C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 881 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-525-940-21

Query Match          9.1%  Score 82.5; DB 2; Length 881;
Best Local Similarity 26.0%  Pred. No. 1.5;
Matches 27; Conservative 10; Mismatches 34; Indels 33; Gaps

Qy 27 TGVAMRSCPESQYWDPLLGTCMSCKTTCNHQSQTCAAF-----CR-----SLSC 71
      | : ||: | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 685 TNSCVTHCPDGSQDTKKNLKRKSENC-----KTCTEFHNCTCRDGLSLQGSRCVSVC 739
      | : ||: | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 72 RKEQGFVDHLLRCISCASCIC-----GOHPKQC-----AVFCEN 106
      | : ||: | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 740 --EGRFYNG--QDCQCHRFCAATCAGAGAGCINCIEGYFME 779
      | : ||: | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 8
US-08-976-838-21
; Sequence 21, Application US/08976838
; Patent No. 5981259
; GENERAL INFORMATION:
; APPLICANT: Franzusoff, Alex
; TITLE OF INVENTION: CD4+ T-LYMPHOCYTE PROTEASE NUCLEIC ACID
; TITLE OF INVENTION: MOLECULES
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross P.C.
; STREET: 1700 Lincoln St., Suite 3500
; CITY: Denver
; STATE: Colorado
; COUNTRY: U.S.A.
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/976,838
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Connell, Gary J.
; REGISTRATION NUMBER: 32,020
; REFERENCE/DOCKET NUMBER: 2848-11-C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 881 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-976-838-21

Query Match          9.1%  Score 82.5; DB 2; Length 881;
Best Local Similarity 26.0%  Pred. No. 1.5;
Matches 27; Conservative 10; Mismatches 34; Indels 33; Gaps

```


; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Ancylostoma caninum
US-08-465-380-21

Query Match 8.7%; Score 79.5; DB 2; Length 98;
Best Local Similarity 28.2%; Pred. No. 0.19;
Matches 22; Conservative 11; Mismatches 34; Indels 11; Gaps 4;
QY 27 TGVAMRSCPE--EQYWDPLLGTCMCKTICNHQSQRTCAAFCSRSLSCRK-----EQGKF 78
Db 19 TRTVRKAYPECGENELWDVCGTKRKPCEAKGSEEEED--PICRSFSCPGPAACVCEGDFY 76

QY 79 YDHLRLDCISCASICGOH 96
Db 77 RDTVIGDCVK-EEECDQH 93

RESULT 14
US-08-480-478-49
; Sequence 49, Application US/08480478
; Patent No. 5864009
; GENERAL INFORMATION:
; APPLICANT: GEORGE P. VLASUK; PATRICK ERIC
; APPLICANT: HUGO STANSSENS; JORIS HILDA
; APPLICANT: LIEVEN MESSENS; MARC JOZEF
; APPLICANT: LAUWERYS; YVES RENE LAROCHE;
; APPLICANT: LAURENT STEPHANE JESPEERS; and
; APPLICANT: YANNICK GEORGES JOZEF
; APPLICANT: GANSEMANS
; TITLE OF INVENTION: NEMATODE-EXTRACTED ANTI-
; TITLE OF INVENTION: COAGULANT PROTEIN
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071

COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: Storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,478
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/326,110
; FILING DATE: 18 OCTOBER 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 208/290
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 98 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-480-478-49

Query Match 8.7%; Score 79.5; DB 2; Length 98;
Best Local Similarity 28.2%; Pred. No. 0.19;
Matches 22; Conservative 11; Mismatches 34; Indels 11; Gaps 4;
QY 27 TGVAMRSCPE--EQYWDPLLGTCMCKTICNHQSQRTCAAFCSRSLSCRK-----EQGKF 78
Db 19 TRTVRKAYPECGENELWDVCGTKRKPCEAKGSEEEED--PICRSFSCPGPAACVCEGDFY 76

QY 79 YDHLRLDCISCASICGOH 96
Db 77 RDTVIGDCVK-EEECDQH 93

RESULT 15
US-08-486-397-21
; Sequence 21, Application US/08486397
; Patent No. 5866542
; GENERAL INFORMATION:
; APPLICANT: George P. Vlasuk, Patric H. Stanssens,
; APPLICANT: Joris H.L. Mensens, Marc J. Lauwereys,
; APPLICANT: Yves R. Laroche, Laurent S. Jespers,
; APPLICANT: Yannick G.J. Gansemans, Matthew Moyle,
; APPLICANT: Peter W. Bergum
; TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 357
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071

COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: Storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,397
; FILING DATE: June 5, 1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/326,110
; FILING DATE: October 18, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 213/269
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 98 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Ancylostoma caninum
US-08-486-397-21

Query Match 8.7%; Score 79.5; DB 2; Length 98;
Best Local Similarity 28.2%; Pred. No. 0.19;
Matches 22; Conservative 11; Mismatches 34; Indels 11; Gaps 4;
QY 27 TGVAMRSCPE--EQYWDPLLGTCMCKTICNHQSQRTCAAFCSRSLSCRK-----EQGKF 78
Db 19 TRTVRKAYPECGENELWDVCGTKRKPCEAKGSEEEED--PICRSFSCPGPAACVCEGDFY 76

QY 79 YDHLRDCISCASICQH 96
| : | :
Db 77 RDTVIGDCVK-EECDQH 93

Search completed: June 25, 2002, 16:12:13
Job time: 53 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 25, 2002, 16:13:25 ; Search time 52.94 Seconds
(without alignments)
121.609 Million cell updates/sec

Title: US-09-854-864-16
Perfect score: 405
Sequence: 1 CPEQYWDPLLGTCMSCKT1.....DCISCASICGQHPKQCA YFC 67

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*

1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	81.5	20.1	1299	2	T43251	furin (EC 3.4.21.7)
2	80	19.8	108	2	H69834	hypothetical prote
3	77	19.0	2664	2	T28626	variant-specific s
4	75	18.5	1548	2	S34583	serine proteinase
5	74.5	18.4	962	2	JC5571	subtilisin-like pr
6	74.5	18.4	969	1	A39490	subtilisin-like pr
7	74.5	18.4	975	2	JC5570	subtilisin-like pr
8	71.5	17.7	773	2	I46059	beta-1 integrin su
9	71.5	17.7	932	2	I52527	PACE4A - mouse (fr
10	71.5	17.7	1680	2	A43434	furin (EC 3.4.21.7
11	71	17.5	108	2	AD0715	conserved hypothet
12	70	17.3	1124	1	I58388	protein-tyrosine k
13	69.5	17.2	798	2	S01659	integrin beta-1 ch
14	69.5	17.2	799	1	IJMSFB	fibronectin recept
15	69	17.0	899	2	G02428	subtilisin-like pr
16	69	17.0	915	2	JC6148	subtilisin-like pr
17	68.5	16.9	501	2	I61512	TNF receptor assoc
18	68	16.8	1339	2	JC4387	epidermal growth f
19	67.5	16.7	146	2	S46368	STIG1 protein - co
20	67.5	16.7	2233	2	T28669	surface protein 51
21	67.5	16.7	3084	1	MMWSA	laminin alpha-1 ch
22	67	16.5	1210	2	A52183	epidermal growth f
23	67	16.5	1751	1	MMH0H	laminin alpha-2 ch
24	66.5	16.4	184	2	S43486	B-cell maturation
25	66.5	16.4	937	2	I53282	gene PACE4 protein
26	66	16.3	248	2	T03869	hypothetical prote
27	66	16.3	1119	2	A88481	protein C16A3.6 [i
28	66	16.3	1122	2	I54237	protein-tyrosine k
29	66	16.3	1123	1	JN0712	protein-tyrosine k

ALIGNMENTS

RESULT 1

T43251
furin (EC 3.4.21.75) - fall armyworm
N:Alternate names: paired basic amino acid cleaving enzyme
C:Species: Spodoptera frugiperda (fall armyworm)
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 11-Jan-2000
C:Accession: T43251
R:Cieplik, M.; Klenk, H.
submitted to the EMBL Data Library, January 1996
A:Description: Cloning and functional characterization of
A:Reference number: Z22368
A:Accession: T43251
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-1299 <CIB>
A:Cross-references: EMBL:Z68888; NID:g167859; PID:e219690; PIDN:CAA93116.1
A:Experimental source: clone Sfurin 6; ovary
C:Function:
A:Description: responsible for the endoproteolytic processing of proproteins with spe
C:Keywords: hydrolase; serine proteinase

Query Match 20.1%; Score 81.5; DB 2; Length 1299;

Best Local Similarity 30.1%; Pred. No. 2.2;

Matches 22; Conservative 5; Mismatches 33; Indels 13; Gaps 3;

QY 1 CPEQYWDPLLGTCMSCKTICNHQSORTCAA----FCCRKEQKGFYDHLRLDCISCASIC 56

Db 1116 CLGSQYDATSGTRSCDASC-----RTCSGPGQFSCTGRPLRIDRLNNQVPCCSER 1170

QY 57 G-----QHPKQCA Y 65

Db 1171 GVTNSTPPDCCCH 1183

RESULT 2

H69834
hypothetical protein yhjO - Bacillus subtilis
C:Species: Bacillus subtilis
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 15-Oct-1999
C:Accession: H69834
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber
C.: Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;
A.: Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari,
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gal
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardino
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mau
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portete
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, M.; Rose, M.; Sadale, Y.; Sato, T.; Scanl
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Se

protein-tyrosine k
epidermal growth f
hypothetical prote
hypothetical prote
epidermal growth f
pyruvate ferredoxi
hypothetical prote
hypothetical prote
integrin beta olig
probable proptotel
kinase-related tra
laminin alpha-2 ch
TIM protein, FHL4
integrin beta-1 ch
variant-specific s
eclosion hormone -

30 66 16.3 1125 1 JH0771
31 66 16.3 1210 1 GQHUE
32 65.5 16.2 202 2 T50635
33 65.5 16.2 210 2 T33697
34 65.5 16.2 843 2 A27131
35 65.5 16.2 1173 2 B97208
36 65 16.0 270 2 T27786
37 65 16.0 572 2 T29880
38 65 16.0 799 2 JC4126
39 65 16.0 915 2 B48225
40 65 16.0 1342 2 A36223
41 65 16.0 3106 1 S53868
42 64.5 15.9 279 2 JG0164
43 64.5 15.9 798 2 A28193
44 64.5 15.9 3006 2 T28625
45 64 15.8 97 2 S34769

A;Cross-references: GDB:131390; OMIM:167405

A;Map position: 15q26-15q26

C;Superfamily: subtilisin-like proteinase PACE4; subtilisin homology

C;Keywords: alternative splicing; hydrolase; serine proteinase

F;150-969/Product: serine proteinase PACE4 #status predicted <SIG>

F;196-434/Domain: subtilisin homology <SET>

F;205,246,420/Active site: Asp, His, Ser #status predicted

Query Match 18.4%; Score 74.5; DB 1; Length 969;

Best Local Similarity 27.8%; Pred. No. 8.4;

Matches 15; Conservative 9; Mismatches 23; Indels 7; Gaps 3;

QY 14 CMCKT-ICNHQSORTCAAFCCRKQKGFYDHLRLDCISC---ASIGQHPKQC 63

Db 764 CLSCRRGFYHHQEMNTCTVTLG---PAGFYADESQKNCCKHPCKVCVDEPEKC 814

RESULT 7

JC3570

subtilisin-like proprotein convertase (EC 3.4.21.-) PACE4 precursor, splice form E-I - H

C;Species: Homo sapiens (man)

C;Date: 23-Sep-1997 #sequence_revision 23-Sep-1997 #text_change 20-Jun-2000

C;Accession: JC3570

R;Morl, K.; Kii, S.; Tsuji, A.; Nagahama, M.; Imamaki, A.; Hayashi, K.; Akamatsu, T.; Na

J. Biochem. 121, 941-948, 1997

A;Title: A novel human PACE4 isoform, PACE4E is an active processing protease containing

A;Reference number: JC3570; MUID:97335942

A;Accession: JC3570

A;Status: nucleic acid sequence not shown

A;Molecule type: mRNA

A;Residues: 1-975 <MOR>

A;Cross-references: DDBJ:DB7993; NID:g2330548; PIDN:BAA21791.1; PID:g2330549

A;Experimental source: brain cerebellum

C;Comment: This enzyme is a processing protease and responsible for processing of various

ch. it is retained intracellularly.

C;Genetics:

A;Gene: GDB:PACE4

A;Cross-references: GDB:131390; OMIM:167405

A;Map position: 15q26-15q26

C;Superfamily: subtilisin-like proteinase PACE4; subtilisin homology

C;Keywords: alternative splicing; glycoprotein; hydrolase; serine proteinase

F;1-62/Domain: signal sequence #status predicted <SIG>

F;63-149/Domain: propeptide #status predicted <PRO>

F;196-434/Domain: subtilisin homology <SET>

F;952-968/Domain: hydrophobic cluster #status predicted <HCL>

F;205,246,347,420/Active site: Asp, His, Asn, Ser #status predicted

F;259/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match

Best Local Similarity 27.8%; Pred. No. 8.4;

Matches 15; Conservative 9; Mismatches 23; Indels 7; Gaps 3;

QY 14 CMCKT-ICNHQSORTCAAFCCRKQKGFYDHLRLDCISC---ASIGQHPKQC 63

Db 764 CLSCRRGFYHHQEMNTCTVTLG---PAGFYADESQKNCCKHPCKVCVDEPEKC 814

RESULT 8

I46059

beta-1 integrin subunit - bovine (fragment)

C;Species: Bos primigenius taurus (cattle)

C;Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 20-Aug-1999

C;Accession: I46059

R;MacLaren, L.A.; Wildeman, A.G.

Biol. Reprod. 53, 153-165, 1995

A;Title: Fibronectin receptors in preimplantation development: cloning, expression, and

A;Reference number: I46059; MUID:95399478

A;Accession: I46059

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1-773 <MAC>

A;Cross-references: EMBL:U10865; NID:g520520; PIDN:AAA80571.1; PID:g520521

C;Superfamily: integrin beta chain; laminin-type EGF-like homology

Query Match 17.7%; Score 71.5; DB 2; Length 773;

Best Local Similarity 32.8%; Pred. No. 14;

Matches 21; Conservative 7; Mismatches 29; Indels 7; Gaps 4;

QY 8 DPLLGT--CMSCK--TICNHQSORTCAAFCCRKQKGFYDHLRLDCISQASICGQHPK--Q 62

Db 568 DCSLGTTSNAVNGICNGRGVCEGA--CKCTDPKFGQPTCEMCQCTQLGVCAEHKECVQ 625

QY 63 CAYF 66

Db 626 CRAF 629

RESULT 9

I52527

PACE4A - mouse (fragment)

C;Species: Mus musculus (house mouse)

C;Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 16-Jul-1999

C;Accession: I52527

R;Hosaka, M.; Murakami, K.; Nakayama, K.

Biomed. Res. 15, 383-390, 1994

A;Title: PACE4A is a ubiquitous endoprotease that has similar but not identical subst

A;Reference number: I52527

A;Accession: I52527

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1-932 <RES>

A;Cross-references: GB:D50060; NID:g769700; PIDN:BAA08777.1; PID:g769701

C;Superfamily: subtilisin-like proteinase PACE4; subtilisin homology

F;172-410/Domain: subtilisin homology <SBT>

Query Match

Best Local Similarity 29.6%; Pred. No. 16;

Matches 16; Conservative 7; Mismatches 24; Indels 7; Gaps 3;

QY 14 CMCKT-ICNHQSORTCAAFCCRKQKGFYDHLRLDCISC---ASIGQHPKQC 63

Db 727 CLSCRRGFYHHQEMNTCTVTLG---PAGFYADESQKNCCKHPCKVCVDEPEKC 777

RESULT 10

A43434

furin (EC 3.4.21.75) 2 - fruit fly (Drosophila melanogaster)

C;Species: Drosophila melanogaster

C;Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 31-Mar-2000

C;Accession: A43434

R;Roebroek, A.J.; Creemers, J.W.; Pauli, I.G.; Kurzik-Dunke, U.; Rentrop, M.; Gateff,

J. Biol. Chem. 267, 17208-17215, 1992

A;Title: Cloning and functional expression of Dfurin2, a subtilisin-like proprotein p

A;Reference number: A43434; MUID:92381036

A;Accession: A43434

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-1680 <ROE>

A;Cross-references: GB:M94375; NID:g157461; PID:g157462

A;Note: sequence extracted from NCBI backbone (NCBIN:111933, NCBIPI:111934)

C;Genetics:

A;Gene: FlyBase:Fur2

A;Cross-references: FlyBase:FBgn0004598

C;Superfamily: subtilisin homology

C;Keywords: hydrolase; serine proteinase; transmembrane protein

F;409-652/Domain: subtilisin homology <SBT>

F;418,457,638/Active site: Asp, His, Ser #status predicted

Query Match

Best Local Similarity 32.1%; Pred. No. 23;

Matches 18; Conservative 6; Mismatches 23; Indels 9; Gaps 3;

A; Title: Coordinate induction of fibronectin, fibronectin receptor, tropomyosin, and

Search completed: June 25, 2002, 16:13:26
Job time: 126 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 25, 2002, 16:11:20 ; Search time 52.94 Seconds
(without alignments)
92.568 Million cell updates/sec

Title: US-09-854-864-6
Perfect score: 284
Sequence: 1 MAGQCSQNEYFDSLHACIP.....LTCQRYCNASVTNSVKGTNA 51
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues
Total number of hits satisfying chosen parameters: 283138
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
Database : PIR_71:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	284	100.0	184	2 S43486	B-cell maturation
2	74.5	26.2	1101	2 T16840	hypothetical prote
3	70.5	24.8	1548	2 S34583	serine proteinase
4	69.5	24.5	5376	2 T42215	zonadhesin - mouse
5	66	23.2	63	2 S07127	chymotrypsin/elast
6	66	23.2	330	2 T25169	hypothetical prote
7	65.5	23.1	1299	2 T43251	furin (EC 3.4.21.7
8	62.5	22.0	999	2 T19275	hypothetical prote
9	62.5	22.0	1513	2 T23681	hypothetical prote
10	61.5	21.7	99	2 S60231	gibberellin-regula
11	61.5	21.7	483	2 T24856	hypothetical prote
12	61.5	21.7	520	2 G88846	protein T12A7.2 [i
13	60.5	21.3	1574	2 T13954	MEGF6 protein - ra
14	60.5	21.3	3034	2 T14119	seven-pass transme
15	59	20.8	758	2 T15577	hypothetical prote
16	59	20.8	1717	1 A45558	epidermal growth f
17	58.5	20.6	2533	2 T28675	alpha-51D immobili
18	58.5	20.6	2533	2 T28674	alpha-51D-immobili
19	58	20.4	1816	1 S68960	laminin alpha-4 ch
20	57.5	20.2	1680	2 A43434	furin (EC 3.4.21.7
21	57	20.1	502	2 T20130	hypothetical prote
22	57	20.1	653	2 G96675	hypothetical prote
23	57	20.1	838	2 T20125	hypothetical prote
24	56.5	19.9	701	2 S62460	hypothetical prote
25	56	19.7	447	2 A96639	protein T1P9.18 [i
26	56	19.7	2476	2 T34022	zonadhesin - pig
27	56	19.7	2824	2 T22759	hypothetical prote
28	55.5	19.5	339	1 KHRTB	cathepsin B (EC 3.
29	55.5	19.5	388	2 T31887	hypothetical prote

hypothetical prote
hypothetical prote
hypothetical prote
S-receptor kinase
hypothetical prote
hypothetical prote
cathepsin B (EC 3.
hypothetical prote
finger protein YJL
hypothetical prote
hypothetical prote
vascular endotheli
membrane glycoprot
protein kinase C (
alpha tectorin - m

ALIGNMENTS

RESULT 1
S43486
B-cell maturation factor - human
N;Alternate names: BCM protein; BCM protein; BEL protein
C;Species: Homo sapiens (man)
C;Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 21-Jul-2000
C;Accession: S43486; S31208; S36661
R;Laabi, Y.; Gras, M.P.; Brouet, J.C.; Berger, R.; Larsen, C.J.; Tsapis, A.
Nucleic Acids Res. 22, 1147-1154, 1994
A;Title: The BCM gene, preferentially expressed during B lymphoid maturation, is bid
A;Reference number: S43486; MUID:94218235
A;Accession: S43486
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-184 <LA>
A;Cross-references: EMBL:Z29574; NID:g471244; PIDN:CAA82690.1; PID:g471245
R;Laabi, Y.; Gras, M.P.; Carbonnel, F.; Brouet, J.C.; Berger, R.; Larsen, C.J.; Tsapi
EMBO J. 11, 3897-3904, 1992
A;Title: A new gene, BCM, on chromosome 16 is fused to the interleukin 2 gene by a t
A;Reference number: S31208; MUID:93010984
A;Accession: S31208
A;Molecule type: mRNA
A;Residues: 1-184 <LA2>
A;Cross-references: EMBL:Z14954; NID:g29407; PIDN:CAA78679.1; PID:g29408
A;Accession: S36661
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 4-184 <LA3>
A;Cross-references: EMBL:Z14955
C;Genetics:
A;Gene: GDB:BCMA
A;Cross-references: GDB:135977; OMIM:109545
A;Map position: 16p13.1-16p13.1
A;Introns: 44/1; 93/1
C;Superfamily: human B-cell maturation factor

Query Match 100.0%; Score 284; DB 2; Length 184;
Best Local Similarity 100.0%; Pred. No. 8.2e-25;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGQCSQNEYFDSLHACIPQLRCSNTPTLCQRYCNASVTNSVKGTNA 51
|||||
Db 4 MAGQCSQNEYFDSLHACIPQLRCSNTPTLCQRYCNASVTNSVKGTNA 54
|||||

RESULT 2
T16840
hypothetical protein T10E10.4 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C;Accession: T16840

C;Keywords: cell adhesion

Query Match
Best Local Similarity 24.5%; Score 69.5; DB 2; Length 5376;
Matches 14; Conservative 5; Mismatches 16; Indels 3; Gaps 2;

QY 4 QCSONEYFDSLHLHACIP-CQLRCSNTP--PLTCORYC 38
||| | : |:| ||| :| :| :|
Db 3299 QCPTNSQFTDCLPSCVFCNSCRCEVTSPVPSSCREGC 3336
||| | : |:| ||| :| :| :|

RESULT 5

S07127
chymotrypsin/elastase inhibitor - common roundworm
C;Species: Ascaris lumbricoideus (common roundworm)
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 24-Jul-1997
C;Accession: S07127
R;Babin, D.R.; Peanasky, R.J.; Goos, S.M.
Arch. Biochem. Biophys. 232, 143-161, 1984
A;Title: The inhibitors of chymotrypsin/elastase from Ascaris lumbricoideus: the pr
A;Reference number: S07127; MUID:84255715
A;Accession: S07127
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-63 <BAB>
C;Superfamily: roundworm trypsin inhibitor

Query Match 23.2%; Score 66; DB 2; Length 63;
Best Local Similarity 34.0%; Pred. No. 0.93;
Matches 17; Conservative 7; Mismatches 18; Indels 8; Gaps 4;

QY 5 CSONEYFDSLHLHACIPCOLRC--SSNTLP-PLTQR-YCNASVTSNYKGTN 50
| || :| :| :| ||| ||| :| :| :|
Db 5 CGPNEVWTE---CTGCMCKGPDPENTPCPLMCRRPSCECGSGRMRTN 50
||| :| :| :| ||| ||| :| :| :|

RESULT 6

T25169
hypothetical protein T23F1.6 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
C;Accession: T25169
R;Wilkinson, J.
submitted to the EMBL Data Library, October 1996
A;Reference number: Z19990
A;Accession: T25169
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-330 <NIL>
A;Cross-references: EMBL:Z81129; PIDN:CAB03405.1; GSPDB:GN00023; CESP:T23F1.6
A;Experimental source: clone T23F1
C;Genetics:
A;Gene: CESP:T23F1.6
A;Map position: 5
A;Introns: 16/3
C;Superfamily: gliadin

Query Match 23.2%; Score 66; DB 2; Length 330;
Best Local Similarity 28.3%; Pred. No. 3.8;
Matches 15; Conservative 8; Mismatches 22; Indels 8; Gaps 2;

QY 5 CSONEYFDSLHLHACIP-----COLRCSNTPPL---TCORYCNASVTSNYKGT 49
| : : | :| :| :| :| :| :| :| :| :|
Db 59 CASQQVQLQTSCMPAQCCSQQCOSNTINTOCPTCOOCCQTSSCPMTST 111
| : : | :| :| :| :| :| :| :| :| :

RESULT 7

T43251
furin (EC 3.4.21.75) - fall armyworm
N;Alternate names: paired basic amino acid cleaving enzyme; proprotein convertase; ser

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 25, 2002, 16:12:15 ; Search time 27.55 Seconds
(without alignments)
71.677 Million cell updates/sec

Title: US-09-854-864-6
Perfect score: 284
Sequence: 1 MAGOCSNEYFDSLHCAP.....LTCORYCNASVTNSVKGCTNA 51

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	284	100.0	184	1	TR17_HUMAN	Q02223 homo sapien
2	70.5	24.8	1877	1	PK5_MOUSE	Q04592 mus musculu
3	69.5	24.5	5376	1	ZAN_MOUSE	O08799 mus musculu
4	66	23.2	63	1	ICB1_ASCSU	P07851 ascaris suu
5	62	21.8	1592	1	SORL_CHICK	Q98930 g sortilin-
6	61.5	21.7	99	1	GAS3_ARATH	P46687 arabidopsis
7	60	21.1	867	1	SSPO_BOVIN	P98167 bos taurus
8	60	21.1	1816	1	LM4_MOUSE	P97927 mus musculu
9	58	20.4	1816	1	LM4_HUMAN	Q16363 homo sapien
10	57.5	20.2	1680	1	FUR2_DROME	P30432 drosophila
11	57	20.1	708	1	YB40_HUMAN	Q9ult0 homo sapien
12	56.5	19.9	701	1	YAE5_SCHPO	Q09807 schizosacch
13	56	19.7	2476	1	ZAN_PIG	Q28983 sus scrofa
14	55.5	19.5	339	1	CATB_RAT	P00787 rattus norv
15	55.5	19.5	3695	1	LM45_HUMAN	O15230 homo sapien
16	55	19.4	339	1	CATB_MOUSE	P10605 mus musculu
17	55	19.4	687	1	VS41_GIALA	P92127 giardia lam
18	55	19.4	758	1	YJ06_YEAST	P39529 saccharomyc
19	55	19.4	760	1	EZ_DROME	P42124 drosophila
20	54.5	19.2	419	1	VEGC_HUMAN	P49767 homo sapien
21	54.5	19.2	592	1	PKCZ_MOUSE	Q02956 mus musculu
22	54	19.0	1980	1	MY9B_RAT	Q63358 rattus norv
23	54	19.0	2114	1	MY9B_MOUSE	Q9qy06 mus musculu
24	54	19.0	2158	1	MY9B_HUMAN	Q13459 homo sapien
25	53.5	18.8	323	1	TNR6_BOVIN	P51867 bos taurus
26	53.5	18.8	325	1	VT2_SFVKA	P25943 Shope fibro
27	53.5	18.8	381	1	P53_CANFA	Q29537 canis fami
28	53.5	18.8	1895	1	YLK3_CAEEL	P41951 caenorhabdi
29	53	18.7	131	1	ALX1_MOUSE	P97430 mus musculu
30	53	18.7	455	1	TRIA_HUMAN	P19438 homo sapien
31	53	18.7	591	1	KPCZ_RABIT	O19111 oryctolagus
32	53	18.7	712	1	ENV_HV2S2	P32536 human immun
33	53	18.7	859	1	ENV_HV2ST	P20872 human immun

RESULT 1	TR17_HUMAN	STANDARD;	PRT;	184 AA.
AC	Q02223;			
DT	01-JUL-1993 (Rel. 26, Created)			
DT	01-JUL-1993 (Rel. 26, Last sequence update)			
DT	01-MAR-2002 (Rel. 41, Last annotation update)			
DE	Tumor necrosis factor receptor superfamily member 17 (B-cell maturation protein).			
GN	TNFRSF17 OR BCMA OR BCM.			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND CHROMOSOMAL TRANSLOCATION.			
RC	TISSUE=Peripheral blood leukocytes, and Lymph node;			
RX	MEDLINE=93010984; PubMed=1396583;			
RA	Laabi Y., Gras M.P., Carbonnel F., Brouet J.C., Berger R.,			
RA	Larsen C.J., Tsapis A.;			
RT	"A new gene, BCM, on chromosome 16 is fused to the interleukin 2 gene by a t(4;16)(q26;p13) translocation in a malignant T cell lymphoma.;"			
RL	EMBO J. 11:3897-3904(1992).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=94218235; PubMed=8165126;			
RA	Laabi Y., Gras M., Brouet J., Berger R., Larsen C., Tsapis A.;			
RT	"The BCMA gene, preferentially expressed during B lymphoid maturation, is bidirectionally transcribed.;"			
RL	Nucleic Acids Res. 22:1147-1154(1994).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=99425270; PubMed=10493829;			
RA	Lofthus B.J., Kim U.-J., Sneddon V.P., Kalush F., Brandon R.,			
RA	Fuhrmann J., Mason T., Crosby M.L., Barnstead M., Cronin L.,			
RA	Deslattes Mays A., Cao Y., Xu R.X., Kang H.-L., Mitchell S.,			
RA	Eichler E.E., Harris P.C., Venter J.C., Adams M.D.;			
RT	"Genome duplications and other features in 12 Mb of human chromosome 16p and 16q.;"			
RL	Genomics 60:295-308(1999).			
RN	[4]			
RP	SEQUENCE FROM N.A., AND VARIANT THR-153.			
RA	Kawasaki A., Tsuchiya N., Fukazawa T., Hashimoto H. Tokunaga K.;			
RT	"New polymorphisms of human BCMA.;"			
RL	Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.			
RN	[5]			
RP	FUNCTION.			
RX	MEDLINE=20363816; PubMed=10903733;			
RA	Hatzoglou A., Rousset J., Bourgeade M.F., Rogier E., Madry C.,			
RA	Inoue J., Devergne O., Tsapis A.;			
RT	"TNF receptor family member BCMA (B cell maturation) associates with TNF receptor-associated factor (TRAF) 1, TRAF2, and TRAF3 and activates NF-kappa B, elk-1, c-Jun N-terminal kinase, and p38 mitogen-activated protein kinase.;"			
RL	J. Immunol. 165:1322-1330(2000).			
RN	[6]			

P25391 homo sapien
Q06670 autographa
Q61001 mus musculus
P16344 radianthus
P29615 drosophila
Q28964 macaca mula
Q64181 cavia porce
P18040 human immun
Q9nx02 homo sapien
P46023 lymanaea sta
P33202 saccharomyc
Q9nj15 branchiosto

ALIGNMENTS

RP FUNCTION.
 RX MEDLINE=20259066; PubMed=10801128;
 RA Gross J.A., Johnston J., Mudri S., Enselman R., Dillon S.R.,
 RA Madden K., Xu W., Parrish-Novak J., Foster D., Lofton-Day C.,
 RA Moore M., Littau A., Grossman A., Haugen H., Foley K., Blumberg H.,
 RA Harrison K., Kindsvogel W., Clegg C.H.,
 RT "TACI and BCMA are receptors for a TNF homologue implicated in B-cell
 autoimmune disease.";
 RL Nature 404:995-999(2000).
 CC -1- FUNCTION: Receptor for TNFSF13B/BLYS/BAFF.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN MATURE B-CELLS.
 CC -1- DISEASE: A FORM OF T-CELL ACUTE LYMPHOBLASTIC LEUKEMIA (T-ALL) IS
 CC CHARACTERIZED BY A CHROMOSOMAL TRANSLOCATION T(4;16)(Q26;P13)
 CC WHICH INVOLVES BCMA AND IL2.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z14954; CAA78679.1; -;
 DR EMBL; Z29575; CAA82691.1; -;
 DR EMBL; Z29574; CAA82690.1; -;
 DR EMBL; U95742; AAB67251.1; -;
 DR EMBL; AB052772; BAB60895.1; -;
 DR PIR; S31208; S31208.
 DR PIR; S31209; S31209.
 DR MIM; 109545; -;
 KW Receptor; Proto-oncogene; Transmembrane; Chromosomal translocation;
 KW Polymorphism.
 KW TRANSNM 55 77 POTENTIAL.
 FT SITE 3 4 BREAKPOINT FOR TRANSLOCATION TO FORM
 FT VARIANT 153 153 A->T.
 FT SEQUENCE 184 AA; 20138 MW; 277AF11E276D932 CRC64;
 SQ /FTID=VAR_012234.
 Query Match 100.0%; Score 284; DB 1; Length 184;
 Best Local Similarity 100.0%; Pred. No. 1.9e-27;
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 MAGQCSQNEYFDLSLHACIPQLRCSSTNPPLTCQRYCNASVTSVKGTNA 51
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 4 MAGQCSQNEYFDLSLHACIPQLRCSSTNPPLTCQRYCNASVTSVKGTNA 54
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 RESULT 2
 ID PKC5_MOUSE STANDARD; PRT: 1877 AA.
 AC Q04592; 062040;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Proprotein convertase subtilisin/kexin type 5 precursor (EC 3.4.21.-)
 DE (Proprotein convertase PC5) (Subtilisin/kexin-like protease PC5)
 DE (Convertase PC5) (PC6) (Subtilisin-like proprotein convertase 6)
 DE (SPC6).
 GN PCSK5.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE OF 330-1877 FROM N.A. (ISOFORM PC5B).
 RC STRAIN=ICR; TISSUE=Intestine;
 RX MEDLINE=93327934; PubMed=8335106;
 RA Nakagawa T., Murakami K., Nakayama K.;
 RT "Identification of an isoform with an extremely large Cys-rich region
 RT of PC6, a Kex2-like processing endoprotease.";

RL FEBS Lett. 327:165-171(1993).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM PC5A).
 RC TISSUP=Brain, and Intestine;
 RX MEDLINE=93224489; PubMed=8468318;
 RA Nakagawa T., Hosaka M., Torii S., Watanabe T., Murakami K.,
 RA Nakayama K.;
 RT "Identification and functional expression of a new member of the
 RT mammalian Kex2-like processing endoprotease family: its striking
 RT structural similarity to PACB4.";
 RL J. Biochem. 113:132-135(1993).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM PC5A).
 RC TISSUP=Adrenal cortex;
 RX MEDLINE=93342056; PubMed=8341687;
 RA Lussan J., Vieau D., Hamelin J., Day R., Chretien M., Seidah N.G.;
 RT "cDNA structure of the mouse and rat subtilisin/kexin-like PC5: a
 RT candidate proprotein convertase expressed in endocrine and
 RT nonendocrine cells.";
 RL Proc. Natl. Acad. Sci. U.S.A. 90:6691-6695(1993).
 RN [4]
 RP PARTIAL SEQUENCE, AND SUBCELLULAR LOCATION.
 RX MEDLINE=97103178; PubMed=8947550;
 RA De Bie I., Marcinkiewicz M., Malide D., Lazure C., Nakayama K.,
 RA Bendayan M., Seidah N.G.;
 RT "The isoforms of proprotein convertase PC5 are sorted to different
 RT subcellular compartments.";
 RL J. Cell Biol. 135:1261-1275(1996).
 RN [5]
 RP DEVELOPMENTAL EXPRESSION.
 RX MEDLINE=96293359; PubMed=8698813;
 RA Constam D.B., Calton M., Robertson E.J.;
 RT "SPC4, SPC6, and the novel protease SPC7 are coexpressed with bone
 RT morphogenetic proteins at distinct sites during embryogenesis.";
 RL J. Cell Biol. 134:181-191(1996).
 RN [6]
 RP DEVELOPMENTAL EXPRESSION.
 RX MEDLINE=97436919; PubMed=9291583;
 RA Rancourt S.L., Rancourt D.E.;
 RT "Murine subtilisin-like proteinase SPC6 is expressed during embryonic
 RT implantation, somitogenesis, and skeletal formation.";
 RL Dev. Genet. 21:75-81(1997).
 CC -1- FUNCTION: LIKELY TO REPRESENT A WIDESPREAD ENDOPEPTASE ACTIVITY
 CC WITHIN THE CONSTITUTIVE AND REGULATED SECRETORY PATHWAY. CAPABLE
 CC OF CLEAVAGE AT THE RX(K/R)R CONSENSUS MOTIF. MAY BE RESPONSIBLE
 CC FOR THE MATURATION OF GASTROINTESTINAL PEPTIDES. MAY BE INVOLVED
 CC IN THE CELLULAR PROLIFERATION OF ADRENAL CORTEX VIA THE ACTIVATION
 CC OF GROWTH FACTORS.
 CC -1- CATALYTIC ACTIVITY: RELEASE OF MATURE PROTEINS FROM THEIR
 CC PROPEPTINS BY CLEAVAGE OF ARG-XAA-YAA-ARG-1-ZAA BONDS, WHERE XAA
 CC CAN BE ANY AMINO ACID AND YAA IS ARG OR LYS.
 CC -1- SUBCELLULAR LOCATION: PC5A IS SECRETED THROUGH THE REGULATED
 CC SECRETORY PATHWAY. PC5B IS A TYPE I MEMBRANE PROTEIN LOCALIZED TO
 CC A PARANUCLEAR POST-GOLGI NETWORK COMPARTMENT IN COMMUNICATION WITH
 CC EARLY ENDOSOMES.
 CC -1- ALTERNATIVE PRODUCTS: AT LEAST 2 ISOFORMS: PC5B/LONG (SHOWN HERE)
 CC AND PC5A/SHORT; ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: PC5A IS EXPRESSED IN MOST TISSUES BUT IS MOST
 CC ABUNDANT IN THE INTESTINE AND ADRENALS. PC5B IS EXPRESSED IN THE
 CC INTESTINE, ADRENALS AND LUNG BUT NOT IN THE BRAIN.
 CC -1- DEVELOPMENTAL STAGE: WEAKLY EXPRESSED THROUGHOUT THE EMBRYO.
 CC EXCEPT IN THE DEVELOPING NERVOUS SYSTEM, THE KIDS AND THE LIVER,
 CC BUT MARKEDLY UPREGULATED AT DISCRETE SITES DURING DEVELOPMENT. AT
 CC E6.5, PROMINENT EXPRESSION OBSERVED IN DIFFERENTIATED DECIDUA. AT
 CC E7.5, INTENSE EXPRESSION IN EXTRAEMBRYONIC ENDODERM, AMNION AND
 CC NASCENT MESODERM. AT 8.5, ABUNDANT EXPRESSION IN SOMITES AND YOLK
 CC SAC FOLLOWED BY A CONFINATION TO DERMATOTOME COMPARTMENT. BETWEEN
 CC E9.5 AND E11.5, ABUNDANT EXPRESSION IN AER (THICKENED ECTODERMAL
 CC CELLS OF LIMB BUDS). AT E12.5, EXPRESSION IN THE LIMBS IS CONFINED
 CC TO THE CONDENSING MESENCHYM SURROUNDING THE CARTILAGE. AT THIS
 CC STAGE, STRONG EXPRESSION ALSO DETECTED IN VERTEBRAL AND FACIAL
 CC CARTILAGE PRIMORDIA AND IN THE MUSCLE OF THE TONGUE. AT E16.5,
 CC ABUNDANT EXPRESSION IN EPITHELIAL CELLS OF THE INTESTINAL VILLI.


```

CC EMBL: U97068; AAC26680.1; -.
DR EMBL: U83190; AAC53125.1; -.
DR MGD; MGI:106656; Zan.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR003645; FOLN.
DR InterPro: IPR000998; MAM.
DR InterPro: IPR002919; TIL.
DR InterPro: IPR003328; TILA.
DR InterPro: IPR001007; WMFC.
DR InterPro: IPR001846; Vwd.
DR Pfam: PF00629; MAM; 3.
DR Pfam: PF01826; TIL; 25.
DR Pfam: PF02345; TILA; 25.
DR Pfam: PF00094; vwd; 4.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00274; FOLN; 11.
DR SMART; SM00137; MAM; 2.
DR SMART; SM00214; WVC; 17.
DR SMART; SM00216; VWD; 4.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 18.
DR PROSITE; PS00740; MAM_1; FALSE_NEG.
DR PROSITE; PS50060; MAM_2; 3.
KW Signal; glycoprotein; Transmembrane; EGF-like domain; Cell adhesion;
KW Repeat.
FT CHAIN 1 17
FT SIGNAL 18 5376
FT DOMAIN 18 5376
FT TRANSMEM 5311 5337
FT DOMAIN 5338 5376
FT DOMAIN 45 210
FT DOMAIN 215 374
FT DOMAIN 377 542
FT DOMAIN 547 1170
FT DOMAIN 1171 1280
FT DOMAIN 1281 1669
FT DOMAIN 1670 2056
FT DOMAIN 2057 2459
FT DOMAIN 2460 2579
FT DOMAIN 2580 2699
FT DOMAIN 2700 2819
FT DOMAIN 2820 2939
FT DOMAIN 2940 3059
FT DOMAIN 3060 3179
FT DOMAIN 3180 3299
FT DOMAIN 3300 3416
FT DOMAIN 3417 3536
FT DOMAIN 3537 3656
FT DOMAIN 3657 3776
FT DOMAIN 3777 3892
FT DOMAIN 3893 4928
FT DOMAIN 4029 4148
FT DOMAIN 4149 4263
FT DOMAIN 4264 4283
FT DOMAIN 4384 4503
FT DOMAIN 4504 4623
FT DOMAIN 4624 4743
FT DOMAIN 4744 4863
FT DOMAIN 4864 5261
FT DOMAIN 5259 5295
FT DISULFID 5263 5274
FT DISULFID 5268 5283
FT DISULFID 5285 5294
FT CARBOHYD 339 339
FT CARBOHYD 499 499
FT CARBOHYD 1216 1216
FT CARBOHYD 1239 1239
FT CARBOHYD 1314 1314
FT CARBOHYD 1814 1814
FT CARBOHYD 1908 1908
FT CARBOHYD 1933 1933

```

```

FT CARBOHYD 2028 2028
FT CARBOHYD 2111 2111
FT CARBOHYD 2142 2142
FT CARBOHYD 2332 2332
FT CARBOHYD 2533 2533
FT CARBOHYD 2575 2575
FT CARBOHYD 2692 2692
FT CARBOHYD 2812 2812
FT CARBOHYD 3052 3052
FT CARBOHYD 3065 3065
FT CARBOHYD 3144 3144
FT CARBOHYD 3172 3172
FT CARBOHYD 3288 3288
FT CARBOHYD 3292 3292
FT CARBOHYD 3782 3782
FT CARBOHYD 4005 4005
FT CARBOHYD 4136 4136
FT CARBOHYD 4243 4243
FT CARBOHYD 4254 4254
FT CARBOHYD 4335 4335
FT CARBOHYD 4376 4376
FT CARBOHYD 4586 4586
FT CARBOHYD 5136 5136
FT CARBOHYD 5252 5252
FT CARBOHYD 5376 5376
SQ SEQUENCE 5376 AA; 579908 MW; 0E44DB77DF2A2620 CRC64;

Query Match 24.5%; Score 69.5; DB 1; Length 5376;
Best Local Similarity 36.8%; Pred. No. 3.1;
Matches 14; Conservative 5; Mismatches 16; Indels 3; Gaps 2;

QY 4 QCSONEYFDSLLHACIP-CQLRCSSNTP--PLTCQRYC 38
   || | | | | | | | | | | | | | | | | | | |
Db 3299 QCPTNSOFTCLPSCVPCSNRCVTPSPVSPSCREGC 3336

RESULT 4
ICE1_ASCSU
ID ICE1_ASCSU STANDARD; PRT; 63 AA.
AC P07851;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Chymotrypsin/elastase isoform 1 (C/E-1 inhibitor).
OS Ascaris suum (Pig roundworm) (Ascaris lumbricoidea).
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea;
OC Ascarididae; Ascaris.
OX NCBI_TaxID=6253;
RN [1]
RP SEQUENCE.
RX MEDLINE=84255715; PubMed=6564898;
RT Babin D.R., Peanasky R.J., Goos S.M.;
RT "The isoforms of chymotrypsin/elastase from Ascaris
RT lumbricoidea: the primary structure."
RL Arch. Biochem. Biophys. 232:143-161(1984).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF COMPLEX WITH ELASTASE.
RX MEDLINE=95006335; PubMed=7922044;
RT Huang K., Strynadka N.C., Bernard V.D., Peanasky R.J., James M.N.;
RT "The molecular structure of the complex of Ascaris
RT chymotrypsin/elastase inhibitor with porcine elastase."
RL Structure 2:679-689(1994).
CC -I- FUNCTION: DEFEND THE ORGANISM AGAINST THE HOST'S PROTEINASES.
CC -I- SIMILARITY: BELONGS TO THE ASCARIS TRYPSIN INHIBITOR FAMILY.
DR PIR; S07127; S07127.
DR PDB; LEA1; 05-APR-99.
DR InterPro; IPR002919; TIL.
DR Pfam; PF01826; TIL; 1.
DR Serine protease inhibitor; 3D-structure.
FT DISULFID 5 38
FT DISULFID 14 33
FT DISULFID 17 29
FT DISULFID 21 60

```

[illegible][illegible]

AC P30432;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Furin-like protease 2 precursor (EC 3.4.21.75) (Furin 2).
 GN Fur2.
 OS Drosophila melanogaster (Fruit fly).
 CC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 CC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 CC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=92381036; PubMed=1512259;
 RX Roebroek A.J.M., Creemers J.W.M., Pauli I.G.L., Kurzik-Dumke U.,
 RA Reintrop M., Gateff E.A.F., Leunissen J.A.M., van de Ven W.J.;
 RT "Cloning and functional expression of dfurin2, a subtilisin-like
 RT proprotein processing enzyme of Drosophila melanogaster with multiple
 RT repeats of a cysteine motif.";
 RL J. Biol. Chem. 267:17208-17215(1992).
 CC -!- FUNCTION: FURIN IS LIKELY TO REPRESENT THE UBIQUITOUS ENDOPROTEASE
 CC ACTIVITY WITHIN CONSTITUTIVE SECRETORY PATHWAYS AND CAPABLE OF
 CC CLEAVAGE AT THE RX(K/R)R CONSENSUS MOTIF (BY SIMILARITY).
 CC -!- CATALYTIC ACTIVITY: Release of mature proteins from their
 CC proproteins by cleavage of Arg-Xaa-Arg-|-zaa bonds, where Xaa
 CC can be any amino acid and Yaa is Arg or Lys. Releases albumin,
 CC complement component C3 and von Willebrand factor from their
 CC respective precursors.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8; ALSO KNOWN AS THE
 CC SUBTILASE FAMILY. FURIN SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M94375; AAA28551.1; -
 DR PIR; A43434; A43434.
 DR HSSP; Q99405; IMPT.
 DR FlyBase; FBgn0004598; Fur2.
 DR InterPro; IPR0005561; EGF-like.
 DR InterPro; IPR002174; Furin-like.
 DR InterPro; IPR002884; P_domain.
 DR InterPro; IPR000209; Peptidase_S8.
 DR Pfam; PF00757; Furin-like; 1.
 DR Pfam; PF01483; P; 1.
 DR Pfam; PF00082; Peptidase_S8; 1.
 DR PRINTS; PD00723; SUBTILISIN.
 DR ProDom; PD000717; P_domain; 1.
 DR SMART; SM00181; EGF; 1.
 DR SMART; SM00261; FU; 10.
 DR PROSITE; PS00136; SUBTILASE_ASP; 1.
 DR PROSITE; PS00137; SUBTILASE_HIS; 1.
 DR PROSITE; PS00138; SUBTILASE_SER; 1.
 DR Hydrolase; Serine protease; Glycoprotein; Signal; Transmembrane;
 KW Multigene family; Zymogen; Repeat.
 KW SIGNAL 1 ? POTENTIAL.
 FT PROPEP 1 ? POTENTIAL.
 FT CHAIN 320 1680 FURIN-LIKE PROTEASE 2.
 FT ACT_SITE 418 418 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 457 457 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 638 638 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT DOMAIN 962 1444 10 X TANDEM REPEATS, CYS-RICH.
 FT REPEAT 962 1007 1.
 FT REPEAT 1008 1057 2.
 FT REPEAT 1058 1104 3.
 FT REPEAT 1105 1153 4.
 FT REPEAT 1154 1205 5.
 FT REPEAT 1206 1254 6.
 FT REPEAT 1255 1299 7.

FT REPEAT 1300 1346 8.
 FT REPEAT 1347 1393 9.
 FT REPEAT 1394 1444 10.
 FT TRANSMEM POTENTIAL.
 FT DOMAIN 1508 1532 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 1533 1680
 FT CARBOHYD 3 3
 FT CARBOHYD 109 109 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 130 130 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 203 203 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 443 443 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 481 481 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 928 928 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1061 1061 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1182 1182 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1275 1275 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1278 1278 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1440 1440 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 1680 AA; 183599 MW; 0A99CE8770A8E293 CRC64;
 Query Match 20.2%; Score 57.5; DB 1; Length 1680;
 Best Local Similarity 34.3%; Pred. No. 29;
 Matches 12; Conservative 5; Mismatches 17; Indels 1; Gaps 1;
 Qy 5 CSQNEYFDSLHACIPQLCRSS-NTPLTLCQRYC 38
 |||::: : ||| ||| | |
 Db 1199 CSEFEYSQVEGQCPCHASGCGSGPADTSC TSC 1233
 RESULT 11
 YB40_HUMAN
 ID YB40_HUMAN STANDARD; PRT; 708 AA.
 AC Q9ULT0;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein KIAA1140 (Fragment).
 GN KIAA1140.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=20039618; PubMed=10574461;
 RA Hirose M., Nagase T., Ishikawa K.-I., Kikuno R., Nomura N.,
 RA Ohara O.;
 RT "Characterization of cDNA clones selected by the GeneMark analysis
 RT from size-fractionated cDNA libraries from human brain.";
 RL DNA Res. 6:329-336(1999).
 CC -!- SIMILARITY: CONTAINS 8 TPR REPEATS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AB032966; BAA86454.1; -
 DR InterPro; IPR001440; TPR.
 DR Pfam; PF00515; TPR; 7.
 DR SMART; SM00028; TPR; 3.
 KW Hypothetical protein; Repeat; TPR repeat.
 FT NON_TER 1 1
 FT REPEAT 27 60 TPR 1.
 FT REPEAT 264 297 TPR 2.
 FT REPEAT 347 381 TPR 3.
 FT REPEAT 383 415 TPR 4.
 FT REPEAT 416 449 TPR 5.
 FT REPEAT 595 628 TPR 6.

```
FT REPEAT 630 662 TPR 7.
FT REPEAT 663 696 TPR 8.
SQ SEQUENCE 708 AA; 79140 MW; 8B93440B522CFC1C CRC64;

Query Match 20.1%; Score 57; DB 1; Length 708;
Best Local Similarity 36.8%; Pred. No. 14;
Matches 14; Conservative 6; Mismatches 16; Indels 2; Gaps 1;

QY 5 CSQNEYFDSLHACIPQLRCSSNTPPLTCQRYCNASV 42
   | : | | | | : | | | | | | | | | | : |
Db 277 CGKSAYAVSLLRECV--KLRPSPDTPLMAAKVCIGSL 312

RESULT 12
YAB5_SCHPO
ID YAB5_SCHPO STANDARD; PRT; 701 AA.
AC Q09807;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Hypothetical 81.8 kDa protein C2G11.05C in chromosome I.
GN SPAC2G11.05C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Badcock K., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U40024; AAC48486.1;
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR000998; MAM.
DR InterPro: IPR002919; TIL.
DR InterPro: IPR003328; TILA.
DR InterPro: IPR001007; VWFC.
DR InterPro: IPR001846; Vwd.
DR Pfam: PF00629; MAM; 2.
DR Pfam: PF01826; TIL; 5.
DR Pfam: PF02345; TILA; 5.
DR Pfam: PF00094; Vwd; 4.
DR SMART: SM00181; EGF; 1.
DR SMART: SM00137; MAM; 1.
DR SMART: SM00214; VWC; 2.
DR SMART: SM00216; VWD; 4.
DR PROSITE: PS00022; EGF_1; 1.
DR PROSITE: PS01186; EGF_2; 4.
DR PROSITE: PS00740; MAM_1; 1.
DR PROSITE: PS50060; MAM_2; 2.
KW Signal; Glycoprotein; Transmembrane; EGF-like domain; Cell adhesion;
KW Repeat.
FT SIGNAL 1 29
FT CHAIN 30 2476
FT ZONADHESIN.
FT EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 2419 2439
FT DOMAIN 2440 2476
FT POTENTIAL.
FT CYTOPLASMIC (POTENTIAL).
FT MAM 1.
FT MAM 2.
FT DOMAIN 147 312
FT DOMAIN 319 687
FT 53 X HEPTAPEPTIDE REPEATS (APPROXIMATE)
FT (MUCIN-LIKE DOMAIN).
FT VWF 1 (PARTIAL).
FT DOMAIN 688 799
FT DOMAIN 800 1184
FT DOMAIN 1185 1573
FT VWF 2.
FT VWF 3.
```



```

FT    DISULFID  1459  1468      BY SIMILARITY.
FT    DISULFID  1471  1481      BY SIMILARITY.
FT    DISULFID  1528  1543      BY SIMILARITY.

Query Match      19.5%; Score 55.5; DB 1; Length 3695;
Best Local Similarity 26.7%; Pred. No. le+02;
Matches 16; Conservative 6; Mismatches 19; Indels 19; Gaps 3;

Qy    4  QCSNEYFD---SLLHACIPQ-----LRCSNTPPLTCQRYCNASVTNS 45
      ::  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db    1844  ECAPGFYRDVKGLELGRCPQCQCHGSDRCLPGSGVCVDCQHNTGAHCE- 1902

```

Search completed: June 25, 2002, 16:22:41
Job time: 626 .sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 25, 2002, 16:13:30 ; Search time 89.98 Seconds
(without alignments)
98.052 Million cell updates/sec

Title: US-09-854-864-6
Perfect score: 284
Sequence: 1 MAGCQSQNEVFDLLHACIP.....LTCQRYCNASVTNSVKGTNA 51

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_19:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_prodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	186	65.5	185	11 088472	088472 mus musculus
2	74.5	26.2	966	5 022378	022378 caenorhabdi
3	72.5	25.5	499	11 088714	088714 mus musculus
4	71.5	25.2	249	11 09DB23	09db23 mus musculus
5	71.5	25.2	249	11 09ET35	09et35 mus musculus
6	71	25.0	532	5 017496	017496 ascaris suu
7	71	25.0	718	5 093107	09bi07 entamoeba h
8	69.5	24.5	5374	11 099ND0	099nd0 mus musculus
9	67.5	23.8	293	4 014836	014836 homo sapien
10	67	23.6	62	5 077419	077419 ascaris suu
11	66	23.2	330	5 018118	018118 caenorhabdi
12	65.5	23.1	548	5 09G045	09gq45 giardia lam
13	65.5	23.1	1299	5 026489	026489 spodoptera
14	63	22.2	175	11 098D0	098d0 mus musculus
15	63	22.2	341	11 09D351	09d351 mus musculus
16	62.5	22.0	999	5 017969	017969 caenorhabdi

17	62.5	22.0	1513	5	017970	017970 caenorhabdi
18	61.5	21.7	353	5	09VW81	09vw81 drosophila
19	61.5	21.7	483	5	022423	022423 caenorhabdi
20	61	21.5	263	15	09WEJ8	09wej8 human immun
21	60.5	21.3	1574	11	088281	088281 ratus norv
22	60.5	21.3	3034	11	035161	035161 mus musculu
23	60	21.1	250	12	09DWZ7	09dwz7 spodoptera
24	60	21.1	262	15	09WDX2	09wdx2 human immun
25	60	21.1	369	5	09VW90	09vw90 drosophila
26	60	21.1	387	13	09PVD4	09pvd4 xenopus lae
27	59.5	21.0	856	10	0945R4	0945r4 hordeum vul
28	59	20.8	263	15	09WDU5	09wdu5 human immun
29	59	20.8	263	15	09WDU8	09wdu8 human immun
30	59	20.8	263	15	09WDV0	09wdv0 human immun
31	59	20.8	263	15	09WDV6	09wdv6 human immun
32	59	20.8	263	15	09WDV8	09wdv8 human immun
33	59	20.8	263	15	09WDW7	09wdw7 human immun
34	59	20.8	263	15	09WDW8	09wdw8 human immun
35	59	20.8	263	15	09WDW9	09wdw9 human immun
36	59	20.8	263	15	09WDX0	09wdx0 human immun
37	59	20.8	263	15	09WDX6	09wdx6 human immun
38	59	20.8	263	15	09WEI8	09wei8 human immun
39	59	20.8	263	15	09WEI9	09wei9 human immun
40	59	20.8	263	15	09WEJ0	09wej0 human immun
41	59	20.8	263	15	09WEJ1	09wej1 human immun
42	59	20.8	263	15	09WEJ2	09wej2 human immun
43	59	20.8	263	15	09WEJ4	09wej4 human immun
44	59	20.8	263	15	09WEJ5	09wej5 human immun
45	59	20.8	263	15	09WEJ7	09wej7 human immun

ALIGNMENTS

RESULT	1
088472	
ID	088472 PRELIMINARY; PRT; 185 AA.
AC	088472;
DT	01-NOV-1998 (TREMBLrel. 08, Created)
DT	01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE	B-CELL MATURATION PROTEIN (TUMOR NECROSIS FACTOR RECEPTOR SUPERFAMILY, MEMBER 17).
DE	TFNRSF17.
GN	TFNRSF17.
OS	Mus musculus (Mouse).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX	NCBI_TaxID=10090;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=BALE/C; TISSUE=SPLEEN;
RA	Madry C., Laabi Y., Callebaut I., Roussel J., Hatzoglou A.,
RA	LeConiat M., Mornon J.-P., Berger R., Tsapis A.; Factor Receptor
RT	"Murine BCMA: a new member of the Tumor Necrosis Factor Receptor
RT	Superfamily.";
RL	Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
RN	[2]
RP	SEQUENCE FROM N.A.
RC	STRAIN=C57BL/6J; TISSUE=COLON;
RX	MEDLINE=21085660; PubMed=11217851;
RA	Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA	Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA	Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamakawa I.,
RA	Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA	Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA	Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA	Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA	Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA	Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA	Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldi M.F.,
RA	Brownstein M.J., Bult C., Fletcher C., Fujita M., Griboldi M.,
RA	Gustincich S., Hill D., Hofmann M., Hume D.A., Kamlya M., Lee N.H.,
RA	Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,

```

RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL: AF061505; AAC23799.1; -.
DR EMBL: AK020247; BAB32038.1; -.
DR MGD: MGI:1343050; Tnfrsf17.
SQ SEQUENCE 185 AA; 20442 MW; 8806352B4FD26A8E CRC64;

Query Match 65.5%; Score 186; DB 11; Length 185;
Best Local Similarity 71.4%; Pred. No. 2.8e-19;
Matches 35; Conservative 4; Mismatches 8; Indels 2; Gaps 1;

Qy 1 MAGQCSQNEYFDSLHACIPQLRCSSNTPPLTCQRYCNASVTSYKGT 49
||| || :||||| ||||| || ||||| ||||| ||||| ||||| |||||
Db 1 MAQCQFHSEYFDSLHACKPCHLRCSN--PPATCQPYCDPSVTSYKGT 47

RESULT 2
ID Q22378 PRELIMINARY; PRT; 966 AA.
AC Q22378;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE HYPOTHETICAL 102.5 KDA PROTEIN.
GN T10E10.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
FT NON_TER 1
FT NON_TER 499
SQ SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Geisel C.;
RT "The sequence of C. elegans cosmid T10E10.";
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: U39644; AAA80360.2; -.
DR HSSP: P10969; IWGT.
DR InterPro: IPR002557; Chitin_binding.
DR InterPro: IPR002899; EB.
DR InterPro: IPR000794; Ketoacyl-synt.
DR InterPro: IPR003571; Snake_toxin.
DR Pfam: PF01607; Chitin_Bind_2; 2.
DR SMART: SM00289; WRI; 12.
DR PROSITE: PS00606; B_KETOACYL_SYNTHASE; UNKNOWN_1.
DR PROSITE: PS00272; SNAKE_TOXIN; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 966 AA; 102460 MW; B565A3CDD25216D9 CRC64;

Query Match 26.2%; Score 74.5; DB 5; Length 966;
Best Local Similarity 36.7%; Pred. No. 0.031;
Matches 18; Conservative 11; Mismatches 15; Indels 5; Gaps 3;

```

```

Qy 4 QCSQNEYFDSLHACIPQLR--CSSNTPPLTCQRYCNASVTSYKGTN 50
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 215 QCSQSTFVNSDLNVCVPLAIONSCDSSTQQPVCs--C-SQVSSSCPPTS 260

RESULT 3
ID O88714 PRELIMINARY; PRT; 499 AA.
AC O88714;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE GASTRIC MUCIN-LIKE PROTEIN (FRAGMENT).
GN GASTRIC MUCIN-LIKE.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
FT NON_TER 1
FT NON_TER 499
SQ SEQUENCE FROM N.A.
RC TISSUE=STOMACH;
RA Tomasetto C., Masson R., Wendling C., Lefebvre O., Chenard M.P.,
RA Rio M.C.;
RT "Identification of interactions between trefoil peptides and members
RT of the mucin protein family using the yeast two-hybrid system.";
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ010752; CAA09343.1; -.
DR HSSP: P56682; 1CCV.
DR InterPro: IPR002919; TIL.
DR InterPro: IPR001846; Vwd.
DR Pfam: PF01826; TIL; 1.
DR Pfam: PF00094; vwd; 1.
DR SMART: SM00216; VWD; 1.
FT NON_TER 1
FT NON_TER 499
SQ SEQUENCE 499 AA; 54190 MW; 04F89EF4F23EE61E CRC64;

Query Match 25.5%; Score 72.5; DB 11; Length 499;
Best Local Similarity 48.3%; Pred. No. 0.03;
Matches 14; Conservative 1; Mismatches 7; Indels 7; Gaps 1;

Qy 5 CSQNEYFDSLHACIPQLRCSSNTPPLT 33
||||||| ||| |||
Db 430 CSQNEYFDHSEGTGVC-----APPTT 451

RESULT 4
ID Q9DBZ3 PRELIMINARY; PRT; 249 AA.
AC Q9DBZ3;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE I200009E08RK PROTEIN.
GN TNFRSF13B OR I200009E08RK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=LUNG;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,

```

QY	1	MAGCQSNEYFDSLLHACIPQQRCSNNTPLTCQRYCNASVTSNYRG	48
	:	: : :	
	:	: : :	
	:	: : :	
D6	137	LCGRCDGSGYFDSSTRTCQRCFNCELCTSTNCFCRSNKILTSSNG	184
RESULT	8		
Q99ND0	ID	Q99ND0	PRELIMINARY; PRT; 5374 AA.
AC	Q99ND0;		
DT	01-JUN-2001	(TREMBlrel. 17, Created)	
DT	01-JUN-2001	(TREMBlrel. 17, Last sequence update)	
DT	01-DEC-2001	(TREMBlrel. 19, Last annotation update)	
DE	ZAN (ZONADHESIN).		
OS	ZAN.		
OS	Mus musculus (Mouse).		

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 25, 2002, 16:11:20 ; Search time 88.08 Seconds
(without alignments)
64.314 Million cell updates/sec

Title: US-09-854-864-6

Perfect score: 284

Sequence: 1 MAGQCSONEYFSLHACIP.....LTCORYCNASVTNSVKGTNA 51

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_032802.*

1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	284	100.0	184	21	Amino acid sequenc
2	284	100.0	184	21	A human BCMA prote
3	284	100.0	184	22	Human BCMA protein
4	284	100.0	184	22	Human B cell matur
5	284	100.0	184	22	Human BAFF recepto
6	284	100.0	184	22	Human B cell matur
7	270	95.1	302	22	Human BCMA-Immunog
8	270	95.1	302	22	Mouse IgG signal/h
9	201.5	71.0	157	22	Human BAFF recepto
10	186	65.5	185	22	Amino acid sequenc
11	186	65.5	185	22	Murine B cell matu

12	71.5	25.2	249	21	AA94006
13	67.5	23.8	166	19	AAW75785
14	67.5	23.8	265	19	AAE09244
15	67.5	23.8	293	19	AAW75783
16	67.5	23.8	293	21	AA36312
17	67.5	23.8	293	21	AA94000
18	67.5	23.8	293	22	AAE09240
19	67.5	23.8	293	22	AA971914
20	64.5	22.7	897	22	ABG06309
21	64.5	22.7	897	22	AA65887
22	64.5	22.7	897	22	AA65888
23	61.5	21.7	353	22	AB71555
24	61	21.5	247	21	AA93998
25	60	21.1	369	22	AB71545
26	60	21.1	1792	21	AA48447
27	60	21.1	1816	21	AA48446
28	59	20.8	87	22	AA41056
29	58.5	20.6	796	22	AB63128
30	58.5	20.6	798	22	AB66495
31	58	20.4	1792	21	AA48443
32	58	20.4	1800	21	AA48445
33	58	20.4	1816	21	AA48442
34	58	20.4	1824	21	AA48444
35	57.5	20.2	225	22	AB71511
36	57.5	20.2	1679	22	AB60498
37	57.5	20.2	1679	22	AB60502
38	57	20.1	337	22	AB63629
39	57	20.1	337	22	AAU38954
40	57	20.1	955	22	AB22836
41	57	20.1	1511	22	AB61693
42	56.5	19.9	2820	22	AB63296
43	56.5	19.9	2957	22	AB62214
44	56	19.7	126	22	ABG01557
45	56	19.7	2476	20	AA667738

ALIGNMENTS

RESULT 1

AA08843
ID AAB08843 standard; peptide; 184 AA.

XX AC AAB08843;

XX DT 02-JAN-2001 (first entry)

XX DE Amino acid sequence of human.

XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT Domain 57..77
XX FT /note= "putative transmembrane domain"

XX PN WO200050633-A1.

XX PD 31-AUG-2000.

XX PF 24-FEB-2000; 2000WO-US04925.

XX PR 24-FEB-1999; 99US-0121485.

XX PA (CEHO) GEN HOSPITAL CORP.

XX PI Seed B, Ting A;

XX DR WPI; 2000-558405/51.

XX

PT Identifying a modulator of gene expression for drug designing, by
 PT contacting a compound library with a cell expressing an anti-cell death
 PT gene and reporter gene, and determining alteration in reporter gene
 PT expression -

XX Claim 32; Fig 7A; 53pp; English.

XX The present sequence represents a BCMA (not defined) polypeptide. BCMA
 CC is a necrosis factor (NF)- κ B activator. The method of the invention is
 CC used to identify compounds which modulate BCMA activity (and thus NF- κ B
 CC activity). The specification describes a method of identifying a
 CC polypeptide which increases gene expression from a promoter. The method
 CC involves contacting a library of with a cell which expresses a
 CC recombinant anti-cell death gene and a reporter gene operably linked to
 CC the promoter, and then determining whether the expression of the
 CC reporter gene is altered as a result of contact with library. The method
 CC is useful for identifying polypeptides which increase or decrease gene
 CC expression from a promoter. The BCMA polypeptide or nucleic acid are
 CC useful for preparing a pharmaceutical composition for treating cancer,
 CC apoptosis, viral infections, inflammatory response, such as rheumatoid
 CC arthritis, inflammatory bowel disease or septic shock. BCMA is useful for
 CC identifying compounds that modulate NF- κ B expression and thus for drug
 CC designing.

SQ Sequence 184 AA;

Query Match 100.0%; Score 284; DB 21; Length 184;
 Best Local Similarity 100.0%; Pred. No. 2.7e-26;
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGCQSQNEYFDSLHACIPQLRCSNTPLTCQRYCNASVTSVKGTA 51
 |
 Db 4 magqcsqneyfDSLHACIPQLRCSNTPLTCQRYCNASVTSVKGta 54

RESULT 2

AA94001
 ID AAY94001 standard; Protein; 184 AA.

AC AAY94001;

XX 20-OCT-2000 (first entry)

DE A human BCMA protein, a B cell protein related to TACI.

KW Human; BR43x2; TACI receptor; extracellular domain; BCMA; B cell protein;
 KW transmembrane activator and CAML-interactor; tumour necrosis factor; TNF;
 KW ztnf4 activity; antibody production; autoimmune disease; amyloidosis;
 KW systemic lupus erythematosus; myasthenia gravis; multiple sclerosis;
 KW rheumatoid arthritis; asthma; bronchitis; emphysema; pyelonephritis;
 KW end stage renal failure; glomerulonephritis; vasculitis; nephritis;
 KW renal neoplasms; multiple myeloma; lymphoma; light chain neuropathy;
 KW immune response; immunosuppression; graft rejection; joint pain;
 KW graft versus host disease; inflammation; swelling; anaemia; septic shock;
 KW insulin dependent diabetes mellitus; Crohn's disease; hypertension;
 KW renal artery stenosis; occlusion; cholesterol; renal emboli.

OS Homo sapiens.

XX WO200040716-A2.

PN 13-JUL-2000.

XX 07-JAN-2000; 2000WO-US00396.

PF 07-JAN-1999; 99US-0226533.

XX (ZYMO) ZYMOGENETICS INC.

XX Gross JA, Xu W, Madden K, Yee DP;

XX WPI; 2000-452538/39.

DR N-PSDB; AAA58559.

XX Inhibiting ztnf4 activity in a mammal, to treat autoimmune diseases,
 PT renal disease, graft versus host disease, and inflammation, comprises
 PT administering a BR43x2, TACI or BCMA extracellular domain polypeptide -
 XX Disclosure; Page 152; 175pp; English.

XX The present sequence represents a human BCMA protein, a B cell protein
 CC related to transmembrane activator and CAML-interactor (TACI) receptor.
 CC TACI is a tumour necrosis factor (TNF) receptor. The extracellular
 CC domains of BR43x2 (an isoform of TACI), TACI or BCMA (a related B cell
 CC protein) receptor contain a cysteine rich domain, and are used for
 CC inhibiting ztnf4 activity. ztnf4 is a TNF ligand. They may also be used
 CC for inhibiting BR43x2, TACI or BCMA receptor-ligand engagement associated
 CC with activated or resting B lymphocytes, effector T-cells, or with
 CC antibody production. The antibody production is associated with an
 CC autoimmune disease selected from systemic lupus erythematosus, myasthenia
 CC gravis, multiple sclerosis and rheumatoid arthritis. The ztnf4 activity
 CC and BR43x2, TACI or BCMA receptor-ligand engagement is associated with
 CC asthma, bronchitis, emphysema, end stage renal failure,
 CC glomerulonephritis, vasculitis, nephritis, pyelonephritis, renal
 CC neoplasms, multiple myelomas, lymphomas, light chain neuropathy,
 CC amyloidosis, moderating immune response, immunosuppression, graft
 CC rejection, graft versus host disease, inflammation, insulin dependent
 CC diabetes mellitus, Crohn's disease, joint pain, swelling, anaemia, or
 CC septic shock. BR43x2, TACI, and BCMA polypeptides, fusions, antibodies,
 CC agonists or antagonists can be used to treat hypertension, renal artery
 CC stenosis, or occlusion, and cholesterol or renal emboli.

SQ Sequence 184 AA;

Query Match 100.0%; Score 284; DB 21; Length 184; \

Best Local Similarity 100.0%; Pred. No. 2.7e-26;

Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGCQSQNEYFDSLHACIPQLRCSNTPLTCQRYCNASVTSVKGTA 51

|
 Db 4 magqcsqneyfDSLHACIPQLRCSNTPLTCQRYCNASVTSVKGta 54

RESULT 3

AAE09241

ID AAE09241 standard; Protein; 184 AA.

XX AAE09241;

XX 19-NOV-2001 (first entry)

DE Human BCMA protein.

KW Human; TNF; tumour necrosis factor; TALL-1; APRIL; TNF receptor;
 KW TNFR; TACI; BCMA; therapy; cancer; leukaemia; myeloma; lymphoma;
 KW autoimmune disease; rheumatoid arthritis; multiple sclerosis;
 KW psoriasis.

XX Homo sapiens.

XX WO200160397-A1.

XX 23-AUG-2001.

XX 28-NOV-2000; 2000WO-US32378.

XX 16-FEB-2000; 2000US-0182938.

PR 22-AUG-2000; 2000US-0226986.

XX (GETH) GENENTECH INC.

XX Ashkenazi AJ, Dodge KH, Grewal I, Klim KJ, Marsters SA, Pitti RM;

PI Yan M,

XX

DR WPI: 2001-541628/60.
 DR N-PSDB; AAD15902.
 XX Inhibiting or neutralizing TALL-1 or APRIL polypeptide biological
 PT activity, for treating autoimmune disorders and cancer, comprises
 PT exposing the cells to TALL-1 or APRIL polypeptide agonists or
 PT antagonists.
 XX
 PS Example 2; Fig 2; 160pp; English.
 CC The invention relates to methods of using one or more agonists or
 CC antagonists to modulate the activity of the members of TNF (tumour
 CC necrosis factor) especially TALL-1, APRIL and TNF receptor (TNFR)
 CC e.g. TACI or BCMA. The method is useful for treating pathological
 CC conditions or diseases associated with increased TALL-1 and APRIL
 CC expression or activity. TALL-1 and APRIL antagonists are used to
 CC block the interaction between APRIL and TALL-1 with TACI or BCMA.
 CC They are useful for treating a mammal suffering from cancer such
 CC as leukaemia, lymphoma, myeloma, cancers of lung and colon and
 CC autoimmune diseases e.g. rheumatoid arthritis, multiple sclerosis,
 CC psoriasis and lupus erythematosus. The present sequence is human
 CC BCMA protein.
 XX Sequence 184 AA;
 SQ

Query Match 100.0%; Score 284; DB 22; Length 184;
 Best Local Similarity 100.0%; Pred. No. 2.7e-26;
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGQCSONEYFDSLHACIPCOLRCSSNPPLTCQRYCNASVTSVKGTNA 51
 ||||||||||||||||||||||||||||||||||||||||||||||||
 Db 4 magqcsqneyfDSLHACIPCLRCSSNPPLTCQRYCNASVTSVKGTNA 54

RESULT 4
 AAE00506
 ID AAE00506 standard; Protein; 184 AA.
 XX
 AC AAE00506;
 DT 31-JUL-2001 (first entry)
 XX
 DE Human B cell maturation protein (BCMA).
 KW Human; A Proliferation Inducing Ligand Receptor; APRIL-R; cytostatic;
 KW gene therapy; cancer; nephrotropic; renal disorder; autoimmune disease;
 KW carcinoma; lung; colon; breast; prostate; Grave's disease; hypertension;
 KW systemic lupus erythematosus; SLE; inflammation; cardiovascular disease;
 KW B-cell lympho-proliferative disorder; BCM; immunosuppressive disease;
 KW organ transplantation; HIV; human immunodeficiency virus; TNF;
 KW tumour necrosis factor; BCMA; B cell maturation protein.
 XX
 OS Homo sapiens.
 XX
 XX WO200124811-A1.
 PN
 PD 12-APR-2001.
 XX
 PF 05-OCT-2000; 2000WO-US27579.
 XX
 XX 06-OCT-1999; 99US-0157933.
 PR
 PR 11-FEB-2000; 2000US-0181807.
 PR 30-JUN-2000; 2000US-0215688.
 XX
 XX (BIOJ) BIOGEN INC.
 PA (APOT-) APOTEC R & D SA.
 XX
 PI Schneider P, Thompson J, Cachero T, Ambrose C, Rennett P;
 DR WPI: 2001-266242/27.
 DR N-PSDB; AAD03844.
 XX

PT Treating a mammal for a condition associated with undesired cell
 PT proliferation such as cancer or carcinoma, comprises administering a
 PT composition comprising A Proliferation Inducing Ligand Receptor
 PT (APRIL-R) antagonist.
 XX Claim 3; Fig 3A; 85pp; English.
 PS
 XX The invention relates to a method of treating a mammal for a condition
 CC associated with undesired cell proliferation such as cancer or
 CC carcinoma. The method involves administering a composition comprising
 CC A Proliferation Inducing Ligand Receptor (APRIL-R) also referred as
 CC B cell maturation protein (BCM or BCMA) antagonist that antagonises the
 CC interaction between APRIL and its cognate receptor(s). This method is
 CC useful for treating undesired cell proliferation such as cancer or
 CC carcinoma e.g. human lung carcinoma, colon carcinoma, breast carcinoma,
 CC prostate carcinoma, and other carcinomas whose proliferation is modulated
 CC by APRIL. It is also useful for treating autoimmune diseases (Grave's
 CC disease, systemic lupus erythematosus-SLE); hypertension, cardiovascular
 CC diseases, renal disorders, B-cell lympho-proliferative disorders,
 CC immunosuppressive diseases, organ transplantation, inflammation and
 CC human immunodeficiency virus (HIV), and for treating, suppressing or
 CC altering an immune response involving a signalling pathway between
 CC APRIL-R and its ligand. APRIL-R DNA is also useful in gene therapy.
 CC The present sequence is human APRIL-R also referred as BCMA or
 CC BCM protein.
 XX Sequence 184 AA;
 SQ

Query Match 100.0%; Score 284; DB 22; Length 184;
 Best Local Similarity 100.0%; Pred. No. 2.7e-26;
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGQCSONEYFDSLHACIPCOLRCSSNPPLTCQRYCNASVTSVKGTNA 51
 ||||||||||||||||||||||||||||||||||||||||||||||||
 Db 4 magqcsqneyfDSLHACIPCLRCSSNPPLTCQRYCNASVTSVKGTNA 54

RESULT 5
 AAB60698
 ID AAB60698 standard; Protein; 184 AA.
 XX
 AC AAB60698;
 DT 22-MAY-2001 (first entry)
 XX
 DE Human BAFF receptor (BAFF-R).
 KW Human BAFF-R; BAFF receptor; TNF family; immunoregulatory agent;
 KW immune-related disorder; B-cell growth inhibitor; BCMA;
 KW B-cell maturation inhibitor; immunoglobulin production inhibitor;
 KW autoimmune disorder; B-cell lymphoproliferative disorder; hypertension;
 KW renal disorder; immunosuppressive disorder; HIV infection;
 KW organ transplantation; antiinflammatory; systemic lupus erythematosus;
 KW autoimmune haemolytic anaemia; Grave's disease; multiple myeloma;
 KW B-cell carcinoma; leukaemia; rapidly progressive glomerulonephritis;
 KW lymphoma; gene therapy; cancer; tumour.
 XX
 OS Homo sapiens.
 XX
 XX WO200112812-A2.
 PN
 PD 22-FEB-2001.
 XX
 XX 16-AUG-2000; 2000WO-US22507.
 PF
 XX 17-AUG-1999; 99US-0149378.
 PR
 PR 11-FEB-2000; 2000US-0181684.
 PR 18-FEB-2000; 2000US-0183536.
 XX
 XX (BIOJ) BIOGEN INC.
 PA (APOT-) APOTEC R & D SA.
 XX

CC human BAFF-R may be used in gene therapy to treat tumours, lymphomas,
 CC autoimmune disorders and inherited B-cell-associated disorders. The
 CC present sequence represents the BAFF-R fusion protein BAFF-R-FC,
 CC comprising a mouse IgG-kappa signal sequence, residues 1-153
 CC of human BAFF-R and a human IgG FC sequence.
 XX
 SQ Sequence 302 AA;

Query Match 95.1%; Score 270; DB 22; Length 302;
 Best Local Similarity 96.0%; Pred. No. 2.2e-24;
 Matches 48; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAGQCSQNEYFDSLHACIPQLRCSSTPPLTCQRYCNASVTSVKGTN 50
 |||||
 Db 27 magqcsqneyfdslhacipqlrccsstppltcqrycnasvtsvkgvd 76

RESULT 9
 AAB60700
 ID AAB60700 standard; Protein: 157 AA.
 XX
 AC AAB60700;

DT 22-MAY-2001 (first entry)

DE Human BAFF receptor (BAFF-R) sequence encoded by A plasmid pJST535.

XX Human BAFF-R; BAFF receptor; TNF family; immunoregulatory agent;
 KW immune-related disorder; B-cell growth inhibitor; BCMA;
 KW B-cell maturation inhibitor; immunoglobulin production inhibitor;
 KW autoimmune disorder; B-cell lymphoproliferative disorder; hypertension;
 KW renal disorder; immunosuppressive disorder; HIV infection;
 KW organ transplantation; antiinflammatory; systemic lupus erythematosus;
 KW autoimmune haemolytic anaemia; Grave's disease; multiple myeloma;
 KW B-cell carcinoma; leukaemia; rapidly progressive glomerulonephritis;
 KW lymphoma; gene therapy; cancer; tumour; plasmid pJST535.

XX Homo sapiens.

XX WO200112812-A2.

XX 22-FEB-2001.

XX 16-AUG-2000; 2000WO-US22507.

XX 17-AUG-1999; 99US-0149378.

XX 11-FEB-2000; 2000US-0181684.

XX 18-FEB-2000; 2000US-0183536.

XX (BIOJ) BIOGEN INC.

PA (APOT-) APOTEC R & D SA.

XX MacKay F, Browning J, Ambrose C, Tschopp J, Schneider P;
 PI Thompson J;

XX WPI; 2001-202866/20.

DR N-PSDB; AAF60000.

XX Inhibiting dendritic cell-induced B-cell growth, maturation and B-cell
 PT lympho-proliferative disorder by administering BAFF-receptor
 PT polypeptide, chimeric molecule comprising receptor or anti-BAFF-R
 PT antibody homolog -

XX Example 1; Fig 3; 59pp; English.

XX The invention relates to the use of a BAFF receptor (BAFF-R, also known
 CC as BCMA) protein, or a BAFF-R fusion protein as an agent for the
 CC treatment of a variety of immune-related disorders. BAFF-R is a member of
 CC the TNF (tumour necrosis factor) family, acting as an immunoregulatory
 CC agent, and also plays a role in the development of hypertension and
 CC related disorders. BAFF-R fusion proteins containing it, and BAFF-R-
 CC specific antibodies can be used for inhibiting B-cell growth, dendritic

CC cell-induced B-cell growth and maturation, and immunoglobulin production,
 CC and in the treatment of autoimmune disorders, B-cell lymphoproliferative
 CC disorders, hypertension and renal disorders. The BAFF-R proteins may also
 CC be used in the treatment of immunosuppressive disorders and HIV
 CC infection, and in patients undergoing organ transplantation. The BAFF-R
 CC proteins or BAFF-R specific antibodies may be used for treating,
 CC suppressing or altering an immune response involving a signalling pathway
 CC between BAFF-R and BAFF, thereby inhibiting inflammation. Since BAFF-R
 CC inhibits B-cell growth and maturation it is useful for treating diseases
 CC such as systemic lupus erythematosus, autoimmune haemolytic anaemia,
 CC Grave's disease, multiple myeloma, B-cell carcinomas, leukaemia, rapidly
 CC progressive glomerulonephritis, and lymphomas. Nucleic acids encoding
 CC human BAFF-R may be used in gene therapy to treat tumours, lymphomas,
 CC autoimmune disorders and inherited B-cell-associated disorders. The
 CC present sequence represents a human BAFF-R protein sequence as encoded
 CC by plasmid pJST535. However, this BAFF-R protein sequence is 27 amino
 CC acids shorter than that given in AAB60698.

XX SQ Sequence 157 AA;

Query Match 71.0%; Score 201.5; DB 22; Length 157;
 Best Local Similarity 82.4%; Pred. No. 1.7e-16;
 Matches 42; Conservative 0; Mismatches 0; Indels 9; Gaps 3;

Qy 1 MAGQCSQNEYFDSLHACIPQLRCSSTPPLTCQRYCNASVTSVKGTTA 51
 |||||
 Db 4 mag--qneyfdslhacipqlr---ntppltcqrycnasvtn---gtta 45

RESULT 10

AAB08844

ID AAB08844 standard; peptide; 185 AA.

XX AAB08844;

DT 02-JAN-2001 (first entry)

XX Amino acid sequence of murine BCMA polypeptide.

XX BCMA: necrosis factor-kB activator; NF-kB; gene expression; cancer;
 KW anti-cell death gene; apoptosis; viral infection; inflammatory response;
 KW rheumatoid arthritis; inflammatory bowel disease; septic shock.

XX Mus musculus.

XX Key Location/Qualifiers

FT Domain 47..72 /note= "putative transmembrane domain"

XX WO200050633-A1.

XX 31-AUG-2000.

XX 24-FEB-2000; 2000WO-US04925.

XX 24-FEB-1999; 99US-0121485.

XX (GEHO) GEN HOSPITAL CORP.

XX Seed B, Ting A;

XX WPI; 2000-558405/51.

XX Identifying a modulator of gene expression for drug designing, by
 PT contacting a compound library with a cell expressing an anti-cell death
 PT gene and reporter gene, and determining alteration in reporter gene
 PT expression -

XX Claim 32; Fig 7B; 53pp; English.

XX The present sequence represents a BCMA (not defined) polypeptide. BCMA
 CC is a necrosis factor (NF)-kB activator. The method of the invention is

used to identify compounds which modulate BCMA activity (and thus NF-kB activity). The specification describes a method of identifying a polypeptide which increases gene expression from a promoter. The method involves contacting a library of with a cell which expresses a recombinant anti-cell death gene and a reporter gene operably linked to the promoter, and then determining whether the expression of the reporter gene is altered as a result of contact with library. The method is useful for identifying polypeptides which increase or decrease gene expression from a promoter. The BCMA polypeptide or nucleic acid are useful for preparing a pharmaceutical composition for treating cancer, apoptosis, viral infections, inflammatory response, such as rheumatoid arthritis, inflammatory bowel disease or septic shock. BCMA is useful for identifying compounds that modulate NF-kB expression and thus for drug designing.

Sequence 185 AA;

Query Match 55.5%; Score 186; DB 21; Length 185;
Best Local Similarity 71.4%; Pred. No. 1.4e-14;
Matches 35; Conservative 4; Mismatches 8; Indels 2; Gaps 1;

QY 1 MAGQCSONEYFDSLHACIPQLRCSNTPLTCORYCNASVTNSVKGT 49
Db 1 maqqcfhsefyfslhacpqlrcsn--ppatcqpcdpbstsvskgt 47

RESULT 11

AAAY71980

ID AAY71980 standard; Protein; 185 AA.

XX AC AAY71980;

XX DT 28-MAR-2001 (first entry)

XX DE Murine B cell maturation factor (BCMA) protein.

XX KW Murine; Tumour Necrosis Factor; TNF; immunosuppressant; TALL-1;
KW Tumour necrosis factor and Apol-related Leucocyte expressed Ligand 1;
KW therapy; autoimmune disorder; rheumatoid arthritis; multiple sclerosis;
KW systemic lupus erythematosus; SLE; insulin dependent diabetes mellitus;
KW thrombocytopenia purpura; acute rheumatic fever; Goodpasture's syndrome;
KW haemolytic anaemia; Grave's disease; myasthenia gravis; BCMA;
KW B cell maturation factor; pemphigus vulgaris; B-lymphocyte proliferation;
KW post-streptococcal glomerulonephritis; polyarteritis nodosa.

XX OS Mus musculus.

XX PN WQ200068378-A1.

XX PD 16-NOV-2000.

XX PF 05-MAY-2000; 2000WO-US12266.

XX PR 06-MAY-1999; 99US-0132892.

XX PR 01-MAY-2000; 2000US-0201012.

XX PA (NAJE-) NAT JEWISH MEDICAL & RES CENT.

XX PI Shu HS;

XX XX WPI; 2001-016094/02.

DR N-PSDB; AAD02130.

XX XX Isolated TALL-1 protein is used to identify compounds that regulate B lymphocyte proliferation, used to treat B lymphocyte associated autoimmune disorders -

XX PS Claim 37; Page 107-108; 112pp; English.

XX XX The present invention relates to Tumour necrosis factor (TNF) and Apol-related Leucocyte expressed Ligand 1 (TALL-1) nucleic acid molecules, proteins (including homologues), and their antibodies. The

CC invention in particular relates to methods for regulating the interaction between TALL-1 and TALL-1 receptors (BCMA referred as B cell maturation factor) to regulate monocyte, macrophage and B lymphocyte mediated immune responses. TALL-1 protein is useful for identifying compounds that regulate B lymphocyte proliferation. It is also useful for treating B lymphocyte associated autoimmune disorders like rheumatoid arthritis, systemic lupus erythematosus (SLE), insulin dependant diabetes mellitus, multiple sclerosis, myasthenia gravis, Grave's disease, autoimmune haemolytic anaemia, autoimmune thrombocytopenia purpura, Goodpasture's syndrome, pemphigus vulgaris, acute rheumatic fever, post-streptococcal glomerulonephritis, or polyarteritis nodosa. The TALL-1 protein and its corresponding nucleic acid sequence are also useful in diagnostic assays.

CC The present sequence is a murine B cell maturation factor (BCMA).
XX BCMA is the receptor for TALL-1 protein.

SQ Sequence 185 AA;

Query Match 55.5%; Score 186; DB 22; Length 185;
Best Local Similarity 71.4%; Pred. No. 1.4e-14;
Matches 35; Conservative 4; Mismatches 8; Indels 2; Gaps 1;

QY 1 MAGQCSONEYFDSLHACIPQLRCSNTPLTCORYCNASVTNSVKGT 49
Db 1 maqqcfhsefyfslhacpqlrcsn--ppatcqpcdpbstsvskgt 47

RESULT 12

AAAY94006

ID AAY94006 standard; Protein; 249 AA.

XX AC AAY94006;

XX DT 20-OCT-2000 (first entry)

XX DE A murine ztnf4, a tumour necrosis factor ligand.

XX KW Human; BR43x2; TACI receptor; extracellular domain; BCMA; B cell protein;
KW transmembrane activator and CAML-interactor; tumour necrosis factor; TNF;
KW ztnf4 activity; antibody production; autoimmune disease; amyloidosis;
KW systemic lupus erythematosus; myasthenia gravis; multiple sclerosis;
KW rheumatoid arthritis; asthma; bronchitis; emphysema; pyelonephritis;
KW end stage renal failure; glomerulonephritis; vasculitis; nephritis;
KW renal neoplasm; multiple myeloma; lymphoma; light chain neuropathy;
KW immune response; immunosuppression; graft rejection; joint pain;
KW graft versus host disease; inflammation; swelling; anaemia; septic shock;
KW insulin dependent diabetes mellitus; Crohn's disease; hypertension;
KW renal artery stenosis; occlusion; cholesterol; renal emboli.

XX OS Mus musculus.

XX PN WQ200040716-A2.

XX PD 13-JUL-2000.

XX PF 07-JAN-2000; 2000WO-US00396.

XX PR 07-JAN-1999; 99US-0226533.

XX PA (ZYMO) ZYMOGENETICS INC.

XX PI Gross JA, Xu W, Madden K, Yee DP;

XX DR WPI; 2000-452538/39.

DR N-PSDB; AAA58566.

XX PT Inhibiting ztnf4 activity in a mammal, to treat autoimmune diseases, renal disease, graft versus host disease, and inflammation, comprises administering a BR43x2, TACI or BCMA extracellular domain polypeptide -

PS Disclosure; Page 163; 175pp; English.

XX XX

```
Best Local Similarity 30.6%; Pred. No. 0.56;
Matches 11; Conservative 9; Mismatches 15; Indels 1; Gaps 1;

Qy 5 CSONEYDSSLHACIPQLRCSSNTPPLTCORYCNA 40
   | : : | | | | | | | | | | | | | | | |
Db 34 CPEQYWDPLLGTCMSCKTICNHQS-QRTCAAFCS 68

RESULT 2
US-09-290-333-6
; Sequence 6, Application US/09290333
; Patent No. 6316222
; GENERAL INFORMATION:
; APPLICANT: Bram, Richard J.
; von Bulow, Gotz
; TITLE OF INVENTION: A LYMPHOCYTE SURFACE RECEPTOR THAT BINDS
; CAML, NUCLEIC ACIDS ENCODING THE SAME AND METHODS OF USE
; THEREOF
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th
; Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/810,572A
; FILING DATE: 28-FEB-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 1340-1-007
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 293 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; US-08-810-572A-2

Query Match 23.8%; Score 67.5; DB 2; Length 293;
Best Local Similarity 30.6%; Pred. No. 1;
Matches 11; Conservative 9; Mismatches 15; Indels 1; Gaps 1;

Qy 5 CSONEYDSSLHACIPQLRCSSNTPPLTCORYCNA 40
   | : : | | | | | | | | | | | | | | | |
Db 34 CPEQYWDPLLGTCMSCKTICNHQS-QRTCAAFCS 68

RESULT 4
US-09-290-333-2
; Sequence 2, Application US/09290333
; Patent No. 6316222
; GENERAL INFORMATION:
; APPLICANT: Bram, Richard J.
; von Bulow, Gotz
; TITLE OF INVENTION: A LYMPHOCYTE SURFACE RECEPTOR THAT BINDS
; CAML, NUCLEIC ACIDS ENCODING THE SAME AND METHODS OF USE
; THEREOF
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th
; Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; US-08-810-572A-2

Best Local Similarity 30.6%; Pred. No. 0.56;
Matches 11; Conservative 9; Mismatches 15; Indels 1; Gaps 1;

Qy 5 CSONEYDSSLHACIPQLRCSSNTPPLTCORYCNA 40
   | : : | | | | | | | | | | | | | | | |
Db 34 CPEQYWDPLLGTCMSCKTICNHQS-QRTCAAFCS 68

RESULT 3
US-08-810-572A-2
; Sequence 2, Application US/08810572A
; Patent No. 5969102
; GENERAL INFORMATION:
; APPLICANT: Bram, Richard J.
```


Db 258 QCGANREFDENTCQCV-CKRTCPRNQPLNPGKCACTES 296

RESULT 10

US-08-510-133A-33
; Sequence 33, Application US/08510133A
; Patent No. 6221839
; GENERAL INFORMATION:
; APPLICANT: Alitalo, Karl
; ; Joukov, Vladimir
; TITLE OF INVENTION: Receptor Ligand
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/510,133A
FILING DATE: 01-Aug-1995

CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Gass, David A.
REGISTRATION NUMBER: 38,153
REFERENCE/DOCKET NUMBER: 28113/32863
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 33:

SEQUENCE CHARACTERISTICS:

LENGTH: 350 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 33:
US-08-510-133A-33

Query Match 19.2%; Score 54.5; DB 4; Length 350;
Best Local Similarity 35.0%; Pred. No. 43;
Matches 14; Conservative 2; Mismatches 21; Indels 3; Gaps 2;

QY 4 QCSQNEFYFDSLHACIPQLRCSNTP--PLTCQRYCNAS 41

Db 258 QCGANREFDENTCQCV-CKRTCPRNQPLNPGKCACTES 296

RESULT 11

US-08-585-895-33
; Sequence 33, Application US/08585895
; Patent No. 6245530
; GENERAL INFORMATION:
; APPLICANT: Alitalo, Karl
; ; Joukov, Vladimir
; TITLE OF INVENTION: Receptor Ligand
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/585,895
FILING DATE:

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:
NAME: Gass, David A.
REGISTRATION NUMBER: 38,153
REFERENCE/DOCKET NUMBER: 28113/33072
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 33:

SEQUENCE CHARACTERISTICS:
LENGTH: 350 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-585-895-33

Query Match 19.2%; Score 54.5; DB 4; Length 350;
Best Local Similarity 35.0%; Pred. No. 43;
Matches 14; Conservative 2; Mismatches 21; Indels 3; Gaps 2;

QY 4 QCSQNEFYFDSLHACIPQLRCSNTP--PLTCQRYCNAS 41

Db 258 QCGANREFDENTCQCV-CKRTCPRNQPLNPGKCACTES 296

RESULT 12

US-08-999-811-2
; Sequence 2, Application US/08999811
; Patent No. 5932540
; GENERAL INFORMATION:

APPLICANT: HU, JING-SHAN
APPLICANT: ROSEN, CRAIG A.
APPLICANT: CAO, LIANG

TITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR 2
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:

ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX
STREET: 1100 NEW YORK AVENUE
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/999,811
FILING DATE: HERewith

CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/207,550
FILING DATE: 8-MAR-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/465,968
FILING DATE: 06-JUN-1995

ATTORNEY/AGENT INFORMATION:

NAME: MARKOWICZ, KAREN R.
REGISTRATION NUMBER: 36,351
REFERENCE/DOCKET NUMBER: 1488.1000004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)371-2600
TELEFAX: (202)371-2540

INFORMATION FOR SEQ ID NO: 2:


```

Query Match      19.2%; Score 54.5; DB 4; Length 419;
Best Local Similarity 35.0%; Pred. No. 52;
Matches 14; Conservative 2; Mismatches 21; Indels 3; Gaps 2;

QY  4  QCQNEYFDSLHACIPCLRCSSNTP--PLTCQRYCNAS 41
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db  327  OCGANRFEDNTCCV-CKRTCPNOLNPGKACETES 365

```

Search completed: June 25, 2002, 16:12:12
Job time: 52 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 25, 2002, 16:13:20 ; Search time 52.94 Seconds
(without alignments)
61.712 Million cell updates/sec

Title: US-09-854-864-7

Perfect score: 201

Sequence: 1 CSQNEYFDSLHACIPQLRCSSNTPTTCQRYC 34

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_71:**

1: piri:**

2: pir2:**

3: pir3:**

4: pir4:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	201	100.0	184	2 S43486	B-cell maturation
2	67.5	33.6	1548	2 S34583	serine proteinase
3	64.5	32.1	5376	2 T42215	zonadhesin - mouse
4	61	30.3	1101	2 T16840	hypothetical prote
5	59	29.4	758	2 T15577	hypothetical prote
6	58	28.9	1299	2 T43251	furin (EC 3.4.21.7
7	57.5	28.6	63	2 S07127	chymotrypsin/elast
8	57.5	28.6	1680	2 A43434	epidermal growth f
9	57	28.4	1717	1 A45558	zonadhesin - pig
10	56	27.9	2476	2 T34022	hypothetical prote
11	55.5	27.6	589	2 T01519	hypothetical prote
12	55	27.4	330	2 T25169	hypothetical prote
13	55	27.4	1513	2 T23681	hypothetical prote
14	54.5	27.1	2155	2 T30197	alpha tectorin - m
15	54	26.9	1980	2 S54307	myosin heavy chain
16	54	26.9	2022	2 A59256	finger protein YUL
17	53.5	26.6	758	2 S46625	MEGF6 protein - ra
18	53	26.4	1574	2 T13954	hypothetical prote
19	53	26.4	2824	2 T22759	hypothetical prote
20	52.5	26.1	118	2 S61051	hypothetical prote
21	52	25.9	255	2 A84544	hypothetical prote
22	52	25.9	294	2 T23682	hypothetical prote
23	52	25.9	547	2 T34318	hypothetical prote
24	52	25.9	1474	2 D88550	hypothetical prote
25	52	25.9	2844	2 S28291	protein ZC84.6 [lm
26	51.5	25.6	63	2 S57816	hypothetical prote
27	51.5	25.6	282	2 S35754	antimicrobial pept
28	51.5	25.6	282	2 I48763	siah-1B protein -
29	51.5	25.6	497	2 T27827	siah-1A protein -
					hypothetical prote

30	51.5	25.6	915	2 T21773	hypothetical prote
31	51.5	25.6	927	2 T21772	hypothetical prote
32	51.5	25.6	999	2 T19275	hypothetical prote
33	51.5	25.6	2120	2 T30243	alpha tectorin - c
34	51	25.4	641	2 E96612	probable transcrip
35	51	25.4	653	2 G96675	hypothetical prote
36	51	25.4	746	2 G84605	hypothetical prote
37	50.5	25.1	610	1 I46001	C4b-binding protei
38	50	24.9	63	2 S08572	chymotrypsin/elast
39	50	24.9	74	2 S10332	ubiquitin / riboso
40	50	24.9	342	2 T30370	hypothetical prote
41	50	24.9	376	2 C81272	probable aminotran
42	50	24.9	455	1 GQHUT1	tumor necrosis fac
43	50	24.9	561	2 T27318	hypothetical prote
44	50	24.9	626	2 T27319	hypothetical prote
45	50	24.9	922	2 T23573	hypothetical prote

ALIGNMENTS

RESULT 1

S43486

B-cell maturation factor - human

N;Alternate names: BCM protein; BCMA protein; BEL protein

C;Species: Homo sapiens (man)

C;Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 21-Jul-2000

C;Accession: S43486; S31208; S36661

R;Laabi, Y.; Gras, M.P.; Brouet, J.C.; Berger, R.; Larsen, C.J.; Tsapis, A.

Nucleic Acids Res. 22, 1147-1154, 1994

A;Title: The BCMA gene, preferentially expressed during B lymphoid maturation, is bid

A;Reference number: S43486; MUID:94218235

A;Accession: S43486

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-184 <LA>

A;Cross-references: EMBL:229574; NID:g471244; PIDN:CAA82690.1; PID:g471245

R;Laabi, Y.; Gras, M.P.; Carbonnel, F.; Brouet, J.C.; Berger, R.; Larsen, C.J.; Tsapi

EMBO J. 11, 3897-3904, 1992

A;Title: A new gene, BCM, on chromosome 16 is fused to the interleukin 2 gene by a t

A;Reference number: S31208; MUID:93010984

A;Accession: S31208

A;Molecule type: mRNA

A;Residues: 1-184 <LA>

A;Cross-references: EMBL:214954; NID:g29407; PIDN:CAA78679.1; PID:g29408

A;Accession: S36661

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 4-184 <LA>

A;Cross-references: EMBL:214955

C;Genetics:

A;Gene: GDB:BCMA

A;Cross-references: GDB:135977; OMIM:109545

A;Map position: 16p13.1-16p13.1

A;Introns: 44/1; 93/1

C;Superfamily: human B-cell maturation factor

Query Match 100.0%; Score 201; DB 2; Length 184;

Best Local Similarity 100.0%; Pred. No. 2.7e-17;

Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSQNEYFDSLHACIPQLRCSSNTPTTCQRYC 34

|||||

Db 8 CSQNEYFDSLHACIPQLRCSSNTPTTCQRYC 41

RESULT 2

S34583

serine proteinase (EC 3.4.21.-) PC6B - mouse

C;Species: Mus musculus (house mouse)

C;Date: 02-Dec-1993 #sequence_revision 10-Nov-1995 #text_change 05-Nov-1999

C;Accession: S34583

		Best Local Similarity	37.1%	Prod. NO. 26	Matches	13: Conservative	5: Mismatches	9: Indels	8: Gaps	2:
QY	1	CSQNEYFDSLHACIPQOLRCS	-----	SNTPPLTC	30					
Db	1150	CSRPLRIDRLNNQVCP	---CS	ERGVNTSPPTDC	1181					

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 25, 2002, 16:22:41 ; Search time 27.55 seconds
(without alignments)
47.785 Million cell updates/sec

Title: US-09-854-864-7

Perfect score: 201

Sequence: 1 CSQNEYFDLLHACIPQLRCSSNTPTTCQRYC 34

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	201	100.0	184	1 TR17_HUMAN	Q02223 homo sapien
2	67.5	33.6	1877	1 PCK5_MOUSE	Q04592 mus musculu
3	64.5	32.1	5376	1 ZAN_MOUSE	O88799 mus musculu
4	60	29.9	867	1 SSPO_BOVIN	P98167 bos taurus
5	57.5	28.6	63	1 ICE1_ASCSU	P07851 ascaris suu
6	57.5	28.6	1680	1 FUR2_DROME	P30432 drosophila
7	56	27.9	2476	1 ZAN_PIG	Q28983 sus scrofa
8	55	27.4	708	1 YB40_HUMAN	O9u10 homo sapien
9	54	26.9	1980	1 MY9B_RAT	O63358 rattus norv
10	54	26.9	2114	1 MY9B_MOUSE	O9qy06 mus musculu
11	54	26.9	2158	1 MY9B_HUMAN	Q13459 homo sapien
12	53.5	26.6	381	1 P53_CANFA	Q29537 canis fami
13	53.5	26.6	758	1 YJ06_YEAST	P39529 saccharomyc
14	53	26.4	131	1 ALK1_MOUSE	P97430 mus musculu
15	52.5	26.1	1592	1 SORL_CHICK	Q98930 g sortilin-
16	51.5	25.6	63	1 AMP2_MIRJA	P25404 mirabilis j
17	50.5	25.1	610	1 C4BP_BOVIN	Q28065 bos taurus
18	50.5	25.1	1013	1 PRML_DROME	P82295 drosophila
19	50	24.9	72	1 RL40_TOBAC	P19379 nicotiana t
20	50	24.9	455	1 TRIA_HUMAN	P19438 homo sapien
21	50	24.9	1696	1 PCK5_BRACL	Q9nj15 brachioosto
22	49	24.4	321	1 MEC3_CABEL	P09088 caenorhabdi
23	49	24.4	350	1 PAX4_HUMAN	O43316 homo sapien
24	49	24.4	461	1 TRIA_PIG	P50555 sus scrofa
25	49	24.4	1483	1 UFD4_YEAST	P33202 saccharomyc
26	49	24.4	1984	1 YL_DROME	P98163 drosophila
27	49	24.4	2531	1 NTCL_MOUSE	Q01705 mus musculu
28	49	24.4	3718	1 LMA5_MOUSE	O61001 mus musculu
29	48.5	24.1	323	1 TNR6_BOVIN	P51867 bos taurus
30	48.5	24.1	608	1 ALB1_SALSA	P21848 salmo salar
31	48.5	24.1	608	1 ALB2_SALSA	Q03156 salmo salar
32	48	23.9	56	1 ITR4_RADMA	P16344 radianthus
33	48	23.9	304	1 TFPI_MACMU	Q28864 macaca mula

ALIGNMENTS

RESULT 1	TR17_HUMAN	STANDARD;	PRT;	184 AA.
ID	TR17_HUMAN			
AC	Q02223;			
DT	01-JUL-1993 (Rel. 26, Created)			
DT	01-JUL-1993 (Rel. 26, Last sequence update)			
DT	01-MAR-2002 (Rel. 41, Last annotation update)			
DE	Tumor necrosis factor receptor superfamily member 17 (B-cell maturation protein).			
GN	TNFRSF17 OR BCMA OR BCM.			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata			Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND CHROMOSOMAL TRANSLOCATION, TISSUE-Peripheral blood leukocytes, and Lymph node.			
RC	MEDLINE=93010984; PubMed=1396583;			
RA	Laabi Y., Gras M.P., Carbonnel F., Brouet J.C., Berger R., Larsen C.J., Tsapis A.;			
RT	"A new gene, BCM, on chromosome 16 is fused to the interleukin 2 gene by a t(4;16)(q26;pl3) translocation in a malignant T cell lymphoma.";			
RL	EMBO J. 11:3897-3904(1992).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	MEDLINE=94218235; PubMed=8165126;			
RA	Laabi Y., Gras M., Brouet J., Berger R., Larsen C., Tsapis A.;			
RT	"The BCMA gene, preferentially expressed during B lymphoid maturation is bidirectionally transcribed.";			
RL	Nucleic Acids Res. 22:1147-1154(1994).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	MEDLINE=99425270; PubMed=10493829;			
RA	Loftus B.J., Kim U.-J., Sneddon V.P., Kalush F., Brandon R., Fuhrmann J., Mason T., Crosby M.L., Barnstead M., Gronin L., Deslattes Ways A., Cao Y., Xu R.X., Kang H.-L., Mitchell S., Eichler E.E., Harris P.C., Venter J.C., Adams M.D.;			
RT	"Genome duplications and other features in 12 Mb of DNA sequence from human chromosome 16p and 16q.";			
RL	Genomics 60:295-308(1999).			
RN	[4]			
RP	SEQUENCE FROM N.A., AND VARIANT THR-153.			
RA	Kawasaki A., Tsuchiya N., Fukazawa T., Hashimoto H., Tokunaga K.;			
RT	"New polymorphisms of human BCMA.";			
RL	Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.			
RN	[5]			
RP	FUNCTION.			
RX	MEDLINE=20363816; PubMed=10903733;			
RA	Hatzoglou A., Kousset J., Bourgeade M.F., Rogier E., Madry C., Inoue J., Devergne O., Tsapis A.;			
RT	"TNF receptor family member BCMA (B cell maturation) associates with TNF receptor-associated factor (TRAF) 1, TRAF2, and TRAF3 and activates NF-kappa B, elk-1, c-Jun N-terminal kinase, and p38 mitogen-activated protein kinase.";			
RL	J. Immunol. 165:1322-1330(2000).			
RN	[6]			

P27966 avian rous-
Q43865 homo sapien
P41990 caenorhabdi
Q21444 caenorhabdi
Q9Y468 homo sapien
P29122 homo sapien
P36992 pleuronecte
P41685 felis silve
P55268 homo sapien
P34015 variola vir
P32127 giardia lam
Q09807 schizosacch

RP FUNCTION.
RX MEDLINE=20259066; PubMed=10801128;
RA Gross J.A., Johnston J., Mudri S., Enselman R., Dillon S.R.,
RA Madden K., Xu W., Parrish-Novak J., Foster D., Lofton-Day C.,
RA Moore M., Littau A., Grossman A., Haugen H., Foley K., Blumberg H.,
RA Harrison K., Kindsvogel W., Clegg C.H.,
RT "TAC1 and BCMA are receptors for a TNF homologue implicated in B-cell
autolymphoma disease.";
RL Nature 404:995-999(2000).
CC -1- FUNCTION: Receptor for TNFSF13B/BLYS/BAFF.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN MATURE B-CELLS.
CC -1- DISEASE: A FORM OF T-CELL ACUTE LYMPHOBLASTIC LEUKEMIA (T-ALL) IS
CC CHARACTERIZED BY A CHROMOSOMAL TRANSLOCATION T(4;16)(Q26;P13)
CC WHICH INVOLVES BCMA AND IL12.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z14954; CAA78679.1; -;
DR EMBL; Z29575; CAA82691.1; -;
DR EMBL; Z29574; CAA82690.1; -;
DR EMBL; U95742; AAB67251.1; -;
DR EMBL; AB052772; BAB60895.1; -;
DR PIR; S31208; S31208.
DR PIR; S31209; S31209.
DR MIM; 109545; -;
KW Receptor; Proto-oncogene; Transmembrane; Chromosomal translocation;
KW Polymorphism.
KW TRANSMEM 55 77 POTENTIAL.
FT SITE 3 4 BREAKPOINT FOR TRANSLOCATION TO FORM
FT VARIANT 153 153 A->T. INTERLEUKIN 2/BCM ONCOGENE.
FT SEQUENCE 184 AA; 20138 MW; 277AF11E2767D932 CRC64;
/FTId=VAR_012234.
Query Match 100.0%; Score 201; DB 1; Length 184;
Best Local Similarity 100.0%; Pred. No. 4.8e-19;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 CSQNEYFDLLHACIPCLRCSSNPPLTCQRYC 34
|||||
Db 8 CSQNEYFDLLHACIPCLRCSSNPPLTCQRYC 41
[1]
RESULT 2
ID PK5_MOUSE STANDARD; PRT; 1877 AA.
AC Q04592; 062040;
DT 01-FEB-1995 (Rel. 31, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Proprotein convertase subtilisin/kexin type 5 precursor (EC 3.4.21.-)
DE (Proprotein convertase PC5) (Subtilisin/kexin-like protease PC5)
DE (Convertase PC5) (PC6) (Subtilisin-like proprotein convertase 6)
DE (SPC6).
GN PCSK5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE OF 330-1877 FROM N.A. (ISOFORM PC5B).
RC STRAIN=ICR; TISSUE=Intestine;
RX MEDLINE=93327934; PubMed=8335106;
RA Nakagawa T., Murakami K., Nakayama K.;
RT "Identification of an isoform with an extremely large Cys-rich region
of PC6, a Kex2-like processing endoprotease.";

RL FEBS Lett. 327:165-171(1993).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM PC5A).
RC TISSUE=Brain, and Intestine;
RX MEDLINE=93224489; PubMed=8468318;
RA Nakagawa T., Hosaka M., Torii S., Watanabe T., Murakami K.,
RA Nakayama K.;
RT "Identification and functional expression of a new member of the
mammalian Kex2-like processing endoprotease family: its striking
structural similarity to PACB4.";
RL J. Biochem. 113:132-135(1993).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM PC5A).
RC TISSUE=Adrenal cortex;
RX MEDLINE=93342056; PubMed=8341687;
RA Lussan J., Vleau D., Hamelin J., Day R., Chretien M., Seidah N.G.;
RT "cDNA structure of the mouse and rat subtilisin/kexin-like PC5: a
candidate proprotein convertase expressed in endocrine and
nonendocrine cells.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:6691-6695(1993).
RN [4]
RP PARTIAL SEQUENCE, AND SUBCELLULAR LOCATION.
RX MEDLINE=97103178; PubMed=8947550;
RA De Bie I., Marcinkiewicz M., Malide D., Lazure C., Nakayama K.,
RA Bendayan M., Seidah N.G.;
RT "The isoforms of proprotein convertase PC5 are sorted to different
subcellular compartments.";
RL J. Cell Biol. 135:1261-1275(1996).
RN [5]
RP DEVELOPMENTAL EXPRESSION.
RX MEDLINE=96293359; PubMed=8698813;
RA Constam D.B., Calton M., Robertson E.J.;
RT "SPC4, SPC6, and the novel protease SPC7 are coexpressed with bone
morphogenetic proteins at distinct sites during embryogenesis.";
RL J. Cell Biol. 134:181-191(1996).
RN [6]
RP DEVELOPMENTAL EXPRESSION.
RX MEDLINE=97436919; PubMed=9291583;
RA Rancourt S.L., Rancourt D.E.;
RT "Murine subtilisin-like proteinase SPC6 is expressed during embryonic
implantation, somitogenesis, and skeletal formation.";
RL Dev. Genet. 21:75-81(1997).
CC -1- FUNCTION: LIKELY TO REPRESENT A WIDESPREAD ENDOPEPTIDASE ACTIVITY
CC WITHIN THE CONSTITUTIVE AND REGULATED SECRETORY PATHWAY. CAPABLE
CC OF CLEAVAGE AT THE RX(K/R)R CONSENSUS MOTIF. MAY BE RESPONSIBLE
CC FOR THE MATURATION OF GASTROINTESTINAL PEPTIDES. MAY BE INVOLVED
CC IN THE CELLULAR PROLIFERATION OF ADRENAL CORTEX VIA THE ACTIVATION
CC OF GROWTH FACTORS.
CC -1- CATALYTIC ACTIVITY: RELEASE OF MATURE PROTEINS FROM THEIR
CC PROPEPTIDES BY CLEAVAGE OF ARG-XAA-YAA-ARG-]-ZAA BONDS, WHERE XAA
CC CAN BE ANY AMINO ACID AND YAA IS ARG OR LYS.
CC -1- SUBCELLULAR LOCATION: PC5A IS SECRETED THROUGH THE REGULATED
CC SECRETORY PATHWAY. PC5B IS A TYPE I MEMBRANE PROTEIN LOCALIZED TO
CC A PARANUCLEAR POST-GOLGI NETWORK COMPARTMENT IN COMMUNICATION WITH
CC EARLY ENDOSOMES.
CC -1- ALTERNATIVE PRODUCTS: AT LEAST 2 ISOFORMS: PC5B/LONG (SHOWN HERE)
CC AND PC5A/SHORT; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: PC5A IS EXPRESSED IN MOST TISSUES BUT IS MOST
CC ABUNDANT IN THE INTESTINE AND ADRENALS. PC5B IS EXPRESSED IN THE
CC INTESTINE, ADRENALS AND LUNG BUT NOT IN THE BRAIN.
CC -1- DEVELOPMENTAL STAGE: WEAKLY EXPRESSED THROUGHOUT THE EMBRYO,
CC EXCEPT IN THE DEVELOPING NERVOUS SYSTEM, THE KIDNEYS AND THE LIVER,
CC BUT MARKEDLY UPREGULATED AT DISCRETE SITES DURING DEVELOPMENT. AT
CC E6.5, PROMINENT EXPRESSION OBSERVED IN DIFFERENTIATING DECIDUA. AT
CC E7.5, INTENSE EXPRESSION IN EXTRAEMBRYONIC ENDODERM, AMNION AND
CC NASCENT MESODERM. AT 8.5, ABUNDANT EXPRESSION IN SOMITES AND YOLK
CC SAC FOLLOWED BY A CONFINATION TO DERMATOTOME COMPARTMENT. BETWEEN
CC E9.5 AND E11.5, ABUNDANT EXPRESSION IN AER (THICKENED ECTODERMAL
CC CELLS OF LIMB BUDS). AT E12.5, EXPRESSION IN THE LIMBS IS CONFINED
CC TO THE CONDENSING MESENCHYM SURROUNDING THE CARTILAGE. AT THIS
CC STAGE, STRONG EXPRESSION ALSO DETECTED IN VERTEBRAL AND FACIAL
CC CARTILAGE PRIMORDIA AND IN THE MUSCLE OF THE TONGUE. AT E16.5,
CC ABUNDANT EXPRESSION IN EPITHELIAL CELLS OF THE INTESTINAL VILLI.

CC CC -1- CATALYTIC ACTIVITY: Release of mature proteins from their
 CC CC propeptides by cleavage of Arg-Xaa-Yaa-Arg-l-Zaa bonds, where Xaa
 CC CC can be any amino acid and Yaa is Arg or Lys. Releases albumin,
 CC CC complement component C3 and von Willebrand factor from their
 CC CC respective precursors.
 CC CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8; ALSO KNOWN AS THE
 CC CC SUBTILASE FAMILY, FURIN SUBFAMILY.
 CC CC -----
 CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC CC the European Bioinformatics Institute. There are no restrictions on its
 CC CC use by non-profit institutions as long as its content is in no way
 CC CC modified and this statement is not removed. Usage by and for commercial
 CC CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC CC or send an email to license@isb-sib.ch).
 CC CC -----
 CC CC EMBL: M94375; AAA28551.1; --
 CC CC PIR: A43434; A43434.
 CC CC HSP: Q99405; IMPT.
 CC CC FlyBase: FBgn004598; Fur2.
 CC CC InterPro: IPR000561; EGF-like.
 CC CC InterPro: IPR002174; Furin-like.
 CC CC InterPro: IPR002884; P_domain.
 CC CC InterPro: IPR000209; Peptidase_S8.
 CC CC Pfam: PF00757; Furin-like; 1.
 CC CC Pfam: PF01483; P; 1.
 CC CC PRINTS: PR00082; Peptidase_S8; 1.
 CC CC PROSITE: PR00723; SUBTILISIN.
 CC CC ProDom: PD000717; P_domain; 1.
 CC CC SMART: SM00181; EGF; 1.
 CC CC SMART: SM00261; FU; 10.
 CC CC PROSITE: PS00136; SUBTILASE_ASP; 1.
 CC CC PROSITE: PS00137; SUBTILASE_SER; 1.
 CC CC PROSITE: PS00138; SUBTILASE_HIS; 1.
 CC CC KW Hydrolase; Serine protease; Glycoprotein; Signal; Transmembrane;
 CC CC Multigene family; Zymogen; Repeat.
 CC CC FT SIGNAL 1 ? POTENTIAL.
 CC CC FT PROPEP ? 319 POTENTIAL.
 CC CC FT CHAIN 320 1680 FURIN-LIKE PROTEASE 2.
 CC CC FT ACT_SITE 418 418 CHARGE RELAY SYSTEM (BY SIMILARITY).
 CC CC FT ACT_SITE 457 457 CHARGE RELAY SYSTEM (BY SIMILARITY).
 CC CC FT ACT_SITE 638 638 CHARGE RELAY SYSTEM (BY SIMILARITY).
 CC CC FT DOMAIN 962 1007 10 X TANDEM REPEATS, CYS-RICH.
 CC CC FT REPEAT 962 1007 1.
 CC CC FT REPEAT 1008 1057 2.
 CC CC FT REPEAT 1058 1104 3.
 CC CC FT REPEAT 1105 1153 4.
 CC CC FT REPEAT 1154 1205 5.
 CC CC FT REPEAT 1206 1254 6.
 CC CC FT REPEAT 1255 1299 7.
 CC CC FT REPEAT 1300 1346 8.
 CC CC FT REPEAT 1347 1393 9.
 CC CC FT REPEAT 1394 1444 10.
 CC CC FT TRANSMEM POTENTIAL.
 CC CC FT DOMAIN 1508 1532 CYTOPLASMIC (POTENTIAL).
 CC CC FT DOMAIN 1533 1680
 CC CC FT CARBOHYD 3 3 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC CC FT CARBOHYD 109 109 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC CC FT CARBOHYD 130 130 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC CC FT CARBOHYD 203 203 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC CC FT CARBOHYD 443 443 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC CC FT CARBOHYD 481 481 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC CC FT CARBOHYD 928 928 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC CC FT CARBOHYD 1061 1061 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC CC FT CARBOHYD 1182 1182 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC CC FT CARBOHYD 1275 1275 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC CC FT CARBOHYD 1278 1278 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC CC FT CARBOHYD 1440 1440 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC CC SEQUENCE 1680 AA; 183599 MW; 0A99CE8770A8E293 CRC64;

Query Match

Best Local Similarity 28.6%; Score 57.5; DB 1; Length 1680;

Matches 12; Conservative 5; Mismatches 17; Indels 1; Gaps 1;

QY 1 CSQNEYFDSLHACIPQLRCSS-NTPLTCQRYC 34
 |||::: : || | | | | | | | | | |
 Db 1199 CSEFEYSQVEGQCRPCGSCGSCNGPADTCTSC 1233

RESULT 7
 ZAN_PIG
 ID ZAN_PIG STANDARD; PRT; 2476 AA.
 AC Q28983;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Zonadhesin precursor.
 GN ZAN.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 ON NCBI_TaxID=9823;
 RX [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 823-830; 859-872; 883-890;
 RP 920-925; 960-967; 1235-1244; 1349-1354; 1518-1532; 1624-1656;
 RP 1658-1667; 1777-1795 AND 1914-1921.
 RC STRAIN=MEISHAN; TISSUE=Testis;
 RX MEDLINE=96064658; PubMed=7592795;
 RA Hardy D.M., Garbers D.L.;
 RT "A sperm membrane protein that binds in a species-specific manner to
 RT the egg extracellular matrix is homologous to von Willebrand
 RT factor".
 RL J. Biol. Chem. 270:26025-26028(1995).
 CC -1- FUNCTION: BINDS IN A SPECIES-SPECIFIC MANNER TO THE ZONA PELLUCIDA
 CC OF THE EGG. MAY BE INVOLVED IN GAMETE RECOGNITION AND/OR
 CC SIGNALING.
 CC -1- SUBUNIT: PROBABLY FORMS COVALENT OLIGOMERS.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN, EXCLUSIVELY ON THE
 CC APICAL REGION OF THE SPERM HEAD (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: IN TESTIS, PRIMARILY IN HAPLOID SPERMATIDS.
 CC NOT IN LUNG, LIVER, HEART, SPLEEN, BRAIN, KIDNEY, EPIDIDYMIS.
 CC -1- DOMAIN: THE MAM DOMAINS PROBABLY MEDIATE SPERM ADHESION TO THE
 CC ZONA PELLUCIDA.
 CC -1- DOMAIN: DURING SPERM MIGRATION THROUGH THE REPRODUCTIVE TRACTS,
 CC THE MUCIN-LIKE DOMAIN MIGHT INHIBIT INAPPROPRIATE TRAPPING OF
 CC SPERMATOZOEA OR PROMOTING ADHESION TO THE OVIDUCTAL ISTHMUS.
 CC -1- DOMAIN: THE VWFD DOMAINS 2 AND 3 MAY MEDIATE COVALENT
 CC OLIGOMERIZATION (BY SIMILARITY TO HUMAN INTESTINAL MUCIN MUC2).
 CC -1- PTM: THE MAM DOMAINS AND THE MUCIN-LIKE DOMAINS ARE MISSING
 CC FROM THE ZONADHESIN THAT BINDS TO THE EGG EXTRACELLULAR MATRIX.
 CC PROCESSING MIGHT OCCUR DURING SPERM MATURATION AND/OR
 CC CAPACITATION.
 CC -1- SIMILARITY: CONTAINS 2 MAM DOMAINS.
 CC -1- SIMILARITY: CONTAINS 4.5 VWFD DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U40024; AAC48486.1; --
 CC InterPro: IPR000561; EGF-like.
 CC InterPro: IPR000998; MAM.
 CC InterPro: IPR002919; TIL.
 CC InterPro: IPR003328; TILA.
 CC InterPro: IPR001007; VWFC.
 CC InterPro: IPR001846; Vwd.
 CC Pfam: PF00629; MAM; 2.
 CC Pfam: PF01826; TIL; 5.
 CC Pfam: PF02345; TILA; 5.
 CC Pfam: PF00094; vwd; 4.
 CC SMART: SM00181; EGF; 1.

CC ALSO ACTS AS A GTPASE ACTIVATING PROTEIN ON RHO.
 CC -!- SUBCELLULAR LOCATION: IN UNDIFFERENTIATED CELLS COLOCALIZES WITH
 CC F-ACTIN IN THE CELL PERIPHERY WHILE IN DIFFERENTIATED CELLS ITS
 CC LOCALIZATION IS CYTOPLASMIC WITH THE HIGHEST LEVELS IN THE
 CC PERINUCLEAR REGION (BY SIMILARITY).
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN TESTIS, LUNG, THYMUS, BRAIN,
 CC LIVER, AND SPLEEN.
 CC -!- SIMILARITY: CONTAINS 1 ZINC-DEPENDENT PHORBOL-ESTER AND DAG
 CC BINDING DOMAIN.
 CC -!- SIMILARITY: CONTAINS 1 GAP DOMAIN.
 CC -!- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
 CC -!- SIMILARITY: CONTAINS 4 IQ DOMAINS.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X77609; CAA54700.1; -.
 CC HSP; P08799; 1MND.
 CC InterPro; IPR002219; DAG_PE-bind.
 CC InterPro; IPR000048; IQ.
 CC InterPro; IPR000159; RA.
 CC InterPro; IPR000198; RhoGAP.
 CC InterPro; IPR001609; myosin_head.
 CC Pfam; PF00130; DAG_PE-bind; 1.
 CC Pfam; PF00612; IQ; 4.
 CC Pfam; PF00063; myosin_head; 2.
 CC Pfam; PF00788; RA; 1.
 CC Pfam; PF00620; RhoGAP; 1.
 CC PRINTS; PR00193; MYOSINHEAVY.
 CC ProDom; PD000355; myosin_head; 2.
 CC SMART; SM00109; C1; 1.
 CC SMART; SM0015; IQ; 3.
 CC SMART; SM00242; MYSC; 1.
 CC SMART; SM00314; RA; 1.
 CC SMART; SM00324; RhoGAP; 1.
 CC PROSITE; PS00479; DAG_PE_BIND_DOM_1; 1.
 CC PROSITE; PS50081; DAG_PE_BIND_DOM_2; 1.
 CC PROSITE; PS50096; IQ; 3.
 CC
 CC Myosin; Repeat; ATP-binding; Calmodulin-binding; Actin-binding;
 CC Cytoskeleton; Coiled coil; GTPase activation; Phorbol-ester binding;
 CC Zinc.
 CC
 CC FT DOMAIN 1 940 HEAD OR MOTOR DOMAIN.
 CC FT DOMAIN 941 1045 NECK OR REGULATORY DOMAIN.
 CC FT DOMAIN 1046 1980 TAIL.
 CC FT DOMAIN 1841 1861 COILED COIL (POTENTIAL).
 CC FT DOMAIN 1918 1948 COILED COIL (POTENTIAL).
 CC FT DOMAIN 845 856 ACTIN-BINDING.
 CC FT DOMAIN 958 978 IQ 1.
 CC FT DOMAIN 981 1001 IQ 2.
 CC FT DOMAIN 1002 1024 IQ 3.
 CC FT DOMAIN 1025 1054 IQ 4.
 CC FT DOMAIN 1593 1641 PHORBOL-ESTER AND DAG BINDING.
 CC FT DOMAIN 1673 1822 GAP DOMAIN.
 CC FT NP_BIND 239 246 ATP (POTENTIAL).
 CC SQ SEQUENCE 1980 AA; 225035 MW; D79FEC4D0FAEC05 CRC64;
 CC
 CC Query Match 26.9%; Score 54; DB 1; Length 1980;
 CC Best Local Similarity 52.9%; Pred. No. 28;
 CC Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
 CC
 CC Qy 16 PCQLRCSNTPPLTCOR 32
 CC ||||| 1: ||| :
 CC Db 1812 PCLLRCPDPSDPLTSMK 1828
 CC
 CC RESULT 10
 CC MY9B_MOUSE

ID MY9B_MOUSE STANDARD; PRT; 2114 AA.
 AC Q9QY06; Q9QY07; Q9QY08; Q9QY09;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Myosin IXb (Unconventional myosin-9b).
 GN MY9B OR MYRS.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND VARIANTS.
 RC STRAIN=C57BL/6;
 RX MEDLINE=20047919; PubMed=10580159;
 RA Grewal P.K., Jones A.-M., Maconochie M., Lemmers R.J.F., Frants R.R.,
 RA Hewitt J.E.;
 RT "Cloning of the murine unconventional myosin gene Myo9b and
 RT identification of alternative splicing.";
 RL Gene 240:389-398(1999).
 CC -!- FUNCTION: MYOSINS ARE ACTIN-BASED MOTOR MOLECULES WITH ATPASE
 CC ACTIVITY. UNCONVENTIONAL MYOSINS SERVE IN INTRACELLULAR MOVEMENTS.
 CC MAY BE INVOLVED IN THE REMODELING OF THE ACTIN CYTOSKELETON. BINDS
 CC ACTIN WITH HIGH AFFINITY BOTH IN THE ABSENCE AND PRESENCE OF ATP
 CC AND ITS MECHANOCHEMICAL ACTIVITY IS INHIBITED BY CALCIUM IONS.
 CC ALSO ACTS AS A GTPASE ACTIVATING PROTEIN ON RHO.
 CC -!- SUBCELLULAR LOCATION: IN UNDIFFERENTIATED CELLS COLOCALIZES WITH
 CC F-ACTIN IN THE CELL PERIPHERY WHILE IN DIFFERENTIATED CELLS ITS
 CC LOCALIZATION IS CYTOPLASMIC WITH THE HIGHEST LEVELS IN THE
 CC PERINUCLEAR REGION (BY SIMILARITY).
 CC -!- ALTERNATIVE PRODUCTS: 3 ISOFORMS; 1 (SHOWN HERE), 2/O AND 3/C ARE
 CC PRODUCED BY ALTERNATIVE SPLICING.
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN THE BRAIN, HEART, MUSCLE AND
 CC INNER EAR.
 CC -!- SIMILARITY: CONTAINS 1 ZINC-DEPENDENT PHORBOL-ESTER AND DAG
 CC BINDING DOMAIN.
 CC -!- SIMILARITY: CONTAINS 1 GAP DOMAIN.
 CC -!- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
 CC -!- SIMILARITY: CONTAINS 4 IQ DOMAINS.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF143687; AAF00122.1; -.
 CC EMBL; AF143685; AAF00120.1; -.
 CC EMBL; AF143686; AAF00121.1; ALT_FRAME.
 CC EMBL; AF143683; AAF00118.1; -.
 CC HSP; P08799; 1MND.
 CC MGD; MGI-106624; Myo9b.
 CC InterPro; IPR002219; DAG_PE-bind.
 CC InterPro; IPR000048; IQ.
 CC InterPro; IPR000159; RA.
 CC InterPro; IPR000198; RhoGAP.
 CC InterPro; IPR001609; myosin_head.
 CC Pfam; PF00063; myosin_head; 2.
 CC Pfam; PF00788; RA; 1.
 CC PRINTS; PR00193; MYOSINHEAVY.
 CC ProDom; PD000355; myosin_head; 2.
 CC SMART; SM00109; C1; 1.
 CC SMART; SM00015; IQ; 4.
 CC SMART; SM00242; MYSC; 1.
 CC SMART; SM00314; RA; 1.
 CC SMART; SM00324; RhoGAP; 1.
 CC PROSITE; PS00479; DAG_PE_BIND_DOM_1; 1.
 CC PROSITE; PS50081; DAG_PE_BIND_DOM_2; 1.
 CC PROSITE; PS50096; IQ; 3.
 CC Myosin; Repeat; ATP-binding; Calmodulin-binding; Actin-binding;
 CC Cytoskeleton; Coiled coil; GTPase activation; Phorbol-ester binding;
 CC

zic: Alternative splicing; Polymorphism.
KW FT DOMAIN 1 939 HEAD OR MOTOR DOMAIN.
FT DOMAIN 940 1044 NECK OR REGULATORY DOMAIN.
FT DOMAIN 1045 1144 TAIL.
FT DOMAIN 1045 1144 COILED COIL (POTENTIAL).
FT DOMAIN 1562 1592 COILED COIL (POTENTIAL).
FT DOMAIN 1839 1859 COILED COIL (POTENTIAL).
FT DOMAIN 1915 1945 COILED COIL (POTENTIAL).
FT DOMAIN 844 855 ACTIN-BINDING.
FT DOMAIN 957 977 IQ 1.
FT DOMAIN 979 1000 IQ 2.
FT DOMAIN 1001 1023 IQ 3.
FT DOMAIN 1024 1053 IQ 4.
FT DOMAIN 1591 1639 PHORBOL-ESTER AND DAG BINDING.
FT DOMAIN 1671 1820 GAP DOMAIN.
FT NP_BIND 239 246 ATP (POTENTIAL).
FT VARSPLIC 740 740 R -> RCTGLDFSESELDVNAFEDIMAFVESR (IN ISOFORM 2).
FT VARSPLIC 1875 1890 MISSING (IN ISOFORM 3).
FT VARSPLIC 1977 2114 GPPAPALPCPIPTLSPLPEAAAPRGRTPTSVTVRVKTPR
RTPIMPANIKLPGLPLHTSWAPALQAVVVKRRPEPA
RRQDOVHSYIAPGADLPQSQTILADHDITLPGTKRRYS
PDYCLPPSSGANG -> E (IN ISOFORM 3).
L -> LEVSPVLPSSL (ONLY IN STRAIN
C57B1/6; CONTAINS AN IN-FRAME 33 BP
IMPERFECT DUPLICATION).
FT VARIANT 2083 2083 A -> P.
FT SEQUENCE 2114 AA; 238832 MW; D774D4B1B2788045 CRC64;
Query Match 26.9%; Score 54; DB 1; Length 2114;
Best Local Similarity 52.9%; Pred. No. 30;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
QY 16 PCOLRCSSNPPLTCOR 32
DB 1810 PCLLRCPDSDPLTSMK 1826
RESULT 11
ID MY9B_HUMAN STANDARD: PRT; 2158 AA.
AC Q13459; Q9UHN0; O75314; Q9NUJ2;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin Ixb (Unconventional myosin-9b).
GN MYO9B OR MYR5
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver, and Small intestine;
RX MEDLINE=97063843; PubMed=8907710;
RA Wirth J.A.; Jensen K.A.; Post P.L.; Bement W.M.; Mooseker M.S.;
RT "Human myosin-IXb, an unconventional myosin with a chimerin-like
rho/rac GTPase-activating protein domain in its tail.";
RL J. Cell Sci. 109:653-661(1996).
RN [2]
RP PARTIAL SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RX MEDLINE=20047919; PubMed=10580159;
RA Grewal P.K.; Jones A.-M.; Maconochie M.; Lemmers R.J.F.; Frants R.R.;
RA Hewitt J.E.;
RT "Cloning of the murine unconventional myosin gene Myo9b and
identification of alternative splicing.";
RL Gene 240:389-398(1999).
RN [3]
RP SEQUENCE OF 1940-2158 FROM N.A. (LONG ISOFORM).
RX MEDLINE=98158729; PubMed=9490638;
RA Post P.L.; Bokoch G.M.; Mooseker M.S.;
RT "Human myosin-IXb is a mechanotchemically active motor and a GAP for
rho.";

J. Cell Sci. 111:941-950(1998).
[4]
RN SEQUENCE OF 1828-2023 FROM N.A. (SHORT ISOFORM).
RC TISSUE=Placenta;
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Ishibashi T., Fujimori K.,
RA Tanai H., Kimata M., Watanabe M., Hiraoka S., Ishii S., Kawai Y.,
RA Saito K., Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K.,
RA Masuho Y., Kanehori K.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: MYOSINS ARE ACTIN-BASED MOTOR MOLECULES WITH ATPASE
CC ACTIVITY. UNCONVENTIONAL MYOSINS SERVE IN INTRACELLULAR MOVEMENTS.
CC MAY BE INVOLVED IN THE REMODELING OF THE ACTIN CYTOSKELETON. BINDS
CC ACTIN WITH HIGH AFFINITY BOTH IN THE ABSENCE AND PRESENCE OF ATP
CC AND ITS MECHANOCHEMICAL ACTIVITY IS INHIBITED BY CALCIUM IONS.
CC ALSO ACTS AS A GTPASE ACTIVATING PROTEIN ON RHO.
CC -!- SUBCELLULAR LOCATION: IN UNDIFFERENTIATED CELLS COLOCALIZES WITH
CC F-ACTIN IN THE CELL PERIPHERY WHILE IN DIFFERENTIATED CELLS ITS
CC LOCALIZATION IS CYTOPLASMIC WITH THE HIGHEST LEVELS IN THE
CC PERINUCLEAR REGION.
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS: A LONG FORM (SHOWN HERE) AND A
CC SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -!- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN PERIPHERAL BLOOD
CC LEUKOCYTES AND AT LOWER LEVELS, IN THYMUS, SPLEEN, TESTIS,
CC PROSTATE, OVARY, BRAIN, SMALL INTESTINE, AND LUNG.
CC -!- SIMILARITY: CONTAINS 1 ZINC-DEPENDENT PHORBOL-ESTER AND DAG
CC BINDING DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 GAP DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -!- SIMILARITY: CONTAINS 4 IQ DOMAINS.
CC -!- CAUTION: THE C-TERMINAL SEQUENCE FROM AA 1917 ONWARDS FROM REF.1
CC WAS PROBABLY A CHIMERA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U42391; AAC50402.1; ALT_SEQ.
CC EMBL; AF143684; AAF00119.1; -.
CC EMBL; AF020267; AAC26597.1; -.
CC EMBL; AK02201; BAA92132.1; ALT_INIT.
CC HSSP; P08799; 1MND.
CC MIM; 602129; -.
CC InterPro: IPR002219; DAG_PE-bind.
CC InterPro: IPR000048; IQ.
CC InterPro: IPR000159; RA.
CC InterPro: IPR000198; RhoGAP.
CC InterPro: IPR001609; myosin_head.
CC Pfam; PF00130; DAG_PE-bind; 1.
CC Pfam; PF00612; IQ; 4.
CC Pfam; PF00063; myosin_head; 2.
CC Pfam; PF00788; RA; 1.
CC Pfam; PF00620; RhoGAP; 1.
CC PRINTS; PR00193; MYOSINHEAVY.
CC ProDom; PD000355; myosin_head; 2.
CC SMART; SM00109; CI; 1.
CC SMART; SM00015; IQ; 3.
CC SMART; SM00242; MYSC; 1.
CC SMART; SM00314; RA; 1.
CC SMART; SM00324; RhoGAP; 1.
CC PROSITE; PS00479; DAG_PE_BIND_DOM_1; 1.
CC PROSITE; PS0081; DAG_PE_BIND_DOM_2; 1.
CC PROSITE; PS50096; IQ; 3.
KW Myosin; Repeat; ATP-binding; Calmodulin-binding; Actin-binding;
KW Cytoskeleton; Coiled coil; GTPase activation; Phorbol-ester binding;
KW Zinc; Alternative splicing.
FT DOMAIN 1 939 HEAD OR MOTOR DOMAIN.
FT DOMAIN 940 1044 NECK OR REGULATORY DOMAIN.

RT and two homologues to chromosome III genes.";
 RL Yeast 10:1235-1249(1994).
 CC -|- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -|- SIMILARITY: CONTAINS 1 ZN(2)-CYS(6), FUNGAL-TYPE BINUCLEAR
 CC CLUSTER DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X77688; CAAS4752.1; -;
 DR EMBL; Z49481; CAA89502.1; -;
 DR PIR; S46625; S46625.
 DR SGD; S0003741; XJL206C.
 DR InterPro: IPR001138; ZN2_Cy6_fungal.
 DR Pfam; PF001172; Zn_c1us; 1.
 DR SMART; SM00066; GAL4; 1.
 DR PROSITE; PS00463; ZN2_Cy6_FUNGAL.1; 1.
 DR PROSITE; PS0048; ZN2_Cy6_FUNGAL.2; 1.
 KW Hypothetical protein; Transcription regulation; DNA-binding;
 KW Nuclear protein; Zinc; Metal-binding.
 FT DNA_BIND 47 73 ZN(2)-CYS(6), FUNGAL-TYPE.
 SQ SEQUENCE 758 AA; 86662 MW; 75DA2ECFF173BE75 CRC64;

Query Match 26.6%; Score 53.5; DB 1; Length 758;
 Best Local Similarity 50.0%; Pred. No. 13;
 Matches 11; Conservative 2; Mismatches 6; Indels 3; Gaps 1;

OY 13 ACIPC---QLRCSSTNPPLTCQ 31
 ||||| :||| | | |
 Db 46 ACTACRKRKVRCSGNIPRLCQ 67

RESULT 14
 ALK1_MOUSE STANDARD; PRT; 131 AA.
 AC P97430; O09081; O09082;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Antileukoproteinase 1 precursor (ALP) (Secretory leukocyte protease
 DE inhibitor).
 GN SLPI.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97191310; PubMed=9039268;
 RA Jin F.-Y., Nathan C.F., Radzic D., Ding A.;
 RT "Secretory leukocyte protease inhibitor: a macrophage product induced
 RT by and antagonistic to bacterial lipopolysaccharide.";
 RL Cell 88:417-426(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97271386; PubMed=9126337;
 RA Zitnik R.J., Zhang J., Kashem M.A., Kohno T., Lyons D.E., Wright C.D.,
 RA Rosen E., Goldberg I., Hayday A.C.;
 RT "The cloning and characterization of a murine secretory leukocyte
 RT protease inhibitor cDNA.";
 RL Biochem. Biophys. Res. Commun. 232:687-697(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6 X CBA; TISSUE=Lung;
 RX MEDLINE=98011992; PubMed=9351627;
 RA Abe T., Tomimaga Y., Kikuchi T., Watanabe A., Satoh K., Watanabe Y.,
 RA Nukiwa T.;

RT "Bacterial pneumonia causes augmented expression of the secretory
 RT leukoprotease inhibitor gene in the murine lung.";
 RL Am. J. Respir. Crit. Care Med. 156:1235-1240(1997).
 CC -|- FUNCTION: ACID-STABLE PROTEINASE INHIBITOR WITH STRONG AFFINITIES
 CC FOR TRYPSIN, CHYMOTRYPSIN, ELASTASE, AND CATHEPSIN G. MAY PREVENT
 CC ELASTASE-MEDIATED DAMAGE TO ORAL AND POSSIBLY OTHER MUCOSAL
 CC TISSUES (BY SIMILARITY).
 CC -|- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN LUNG, SPLEEN,
 CC INTESINE AND EPIDIDYMIS WITH LOWER LEVELS IN LIVER AND SEMINAL
 CC VESICLE. NO EXPRESSION IN BRAIN, HEART, KIDNEY AND MUSCLE.
 CC -|- SIMILARITY: CONTAINS 2 WAP-TYPE DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U73004; AAC53047.1; -;
 DR EMBL; U88093; AAC53140.1; -;
 DR EMBL; U94341; AAC53394.1; -;
 DR HSSP; P19957; 1PLE.
 DR MGD; MGI:109297; Slpi.
 DR InterPro: IPR002221; WAP.
 DR Pfam; PF00095; wap; 2.
 DR PRODOM; PD001224; WAP; 1.
 DR SMART; SM00217; WAP; 2.
 DR PROSITE; PS00317; 4_DISULFIDE_CORE; 2.
 KW Serine protease inhibitor; Repeat; Signal.
 FT SIGNAL 1 25
 FT CHAIN 26 131 ANTILEUKOPROTEINASE 1.
 FT DOMAIN 26 84 TRYPSIN INHIBITORY DOMAIN.
 FT DOMAIN 85 131 ELASTASE INHIBITORY DOMAIN.
 FT DOMAIN 32 77 WAP 1.
 FT DOMAIN 86 131 WAP 2.
 FT ACT_SITE 46 46 TRYPSIN INHIBITORY SITE (PROBABLE).
 FT ACT_SITE 98 99 ELASTASE OR CHYMOTRYPSIN INHIBITORY SITE
 FT (PROBABLE).
 FT DISULFID 36 65 BY SIMILARITY.
 FT DISULFID 44 69 BY SIMILARITY.
 FT DISULFID 52 64 BY SIMILARITY.
 FT DISULFID 58 73 BY SIMILARITY.
 FT DISULFID 90 119 BY SIMILARITY.
 FT DISULFID 97 123 BY SIMILARITY.
 FT DISULFID 106 118 BY SIMILARITY.
 FT DISULFID 112 127 BY SIMILARITY.
 SQ SEQUENCE 131 AA; 14308 MW; A57C9E30FE711B8F CRC64;
 Query Match 26.4%; Score 53; DB 1; Length 131;
 Best Local Similarity 47.4%; Pred. No. 3;
 Matches 9; Conservative 1; Mismatches 9; Indels 0; Gaps 0;
 OY 14 CIPQLRCSSTNPPLTCQ 32
 |: ||| |||
 Db 90 CVKQARCMMLNPPNVCQR 108
 RESULT 15
 SORL_CHICK STANDARD; PRT; 1592 AA.
 AC Q98930;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Sortilin-related receptor (Sorting protein-related
 DE LDLR class A repeats) (SORLA) (SoriA-1) (Low-density
 DE receptor relative with 11 ligand-binding repeats) (LDLR relative with
 DE 11 ligand-binding repeats) (LR11) (Fragment).
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

CC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OX Gallus.
 RN NCBI_TaxID=9031;
 [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=97301565; PubMed=9157966;
 RA Morwald S., Yamazaki H., Bujo H., Kusunoki J., Kanaki T., Seimiya K.,
 RA Morisaki N., Nimpf J., Schneider W.J., Saito Y.;
 RA "A novel mosaic protein containing LDL receptor elements is highly
 RT conserved in humans and chickens";
 RL Arterioscler. Thromb. Vasc. Biol. 17:996-1002(1997).
 CC -!- FUNCTION: LIKELY TO BE A MULTIFUNCTIONAL ENDOCYTIC RECEPTOR, THAT
 CC MAY BE IMPLICATED IN THE UPTAKE OF LIPOPROTEINS AND OF PROTEASES.
 CC BINDS LDL, THE MAJOR CHOLESTEROL-CARRYING LIPOPROTEIN OF PLASMA,
 CC AND TRANSPORTS IT INTO CELLS BY ENDOCYTOSIS. BINDS THE RECEPTOR-
 CC ASSOCIATED PROTEIN (RAP). COULD PLAY A ROLE IN CELL-CELL
 CC INTERACTION. BOTH ESTROGEN STATUS AND CHOLESTEROL LEVELS LACK
 CC REGULATORY EFFECTS ON THIS RECEPTOR.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
 CC -!- TISSUE SPECIFICITY: EXPRESSED ABUNDANTLY IN BRAIN. PRESENT IN THE
 CC TESTES, ADRENAL GLANDS AND DETECTABLE IN THE LUNG.
 CC -!- SIMILARITY: CONTAINS 5 BNR REPEATS.
 CC -!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
 CC -!- SIMILARITY: CONTAINS 11 LDL-RECEPTOR CLASS A DOMAINS.
 CC -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Y08109; CAA69324.1; --
 DR HSSP; P01130; 1AJJ.
 DR InterPro; IPR002860; BNR.
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR003961; FN.III.
 DR InterPro; IPR002172; LDLrecept.A.
 DR InterPro; IPR000033; Ldl_receptor_rep.
 DR Pfam; PF02012; BNR; 5.
 DR Pfam; PF00041; fn3; 1.
 DR Pfam; PF00057; ldl_recept_a; 11.
 DR Pfam; PF00058; ldl_recept_b; 5.
 DR PRINTS; PR00261; LDLRECEPTOR.
 DR SMART; SM00181; EGF; 1.
 DR SMART; SM00060; FN3; 1.
 DR SMART; SM00192; LDLa; 11.
 DR SMART; SM00135; LY; 5.
 DR PROSITE; PS01186; EGF_2; 1.
 DR PROSITE; PS01209; LDLRA_1; 11.
 DR PROSITE; PS00068; LDLRA_2; 11.
 KW Endocytosis; Receptor; EGF-like domain; Repeat; Glycoprotein; LDL;
 KW Lipid transport; Cholesterol metabolism.
 FT NON_TER 1 1
 FT REPEAT 43 54 BNR 1.
 FT REPEAT 139 150 BNR 2.
 FT REPEAT 348 359 BNR 3.
 FT REPEAT 428 439 BNR 4.
 FT REPEAT 469 480 BNR 5.
 FT DOMAIN 710 884 5 X APPROXIMATE YWTD REPEATS.
 FT REPEAT 710 713 1.
 FT REPEAT 754 757 2.
 FT REPEAT 798 801 3.
 FT REPEAT 841 844 4.
 FT REPEAT 881 884 5.
 FT DOMAIN 933 979 EGF-LIKE.
 FT DOMAIN 983 1021 LDL-RECEPTOR CLASS A 1.
 FT DOMAIN 1022 1062 LDL-RECEPTOR CLASS A 2.
 FT DOMAIN 1063 1100 LDL-RECEPTOR CLASS A 3.
 FT DOMAIN 1103 1143 LDL-RECEPTOR CLASS A 4.

FT	DOMAIN	1143	1180	LDL-RECEPTOR CLASS A 5.
FT	DOMAIN	1180	1224	LDL-RECEPTOR CLASS A 6.
FT	DOMAIN	1230	1268	LDL-RECEPTOR CLASS A 7.
FT	DOMAIN	1273	1312	LDL-RECEPTOR CLASS A 8.
FT	DOMAIN	1324	1362	LDL-RECEPTOR CLASS A 9.
FT	DOMAIN	1376	1415	LDL-RECEPTOR CLASS A 10.
FT	DOMAIN	1419	1457	LDL-RECEPTOR CLASS A 11.
FT	DOMAIN	1462	1551	FIBRONECTIN TYPE-III 1.
FT	DOMAIN	1559	>1592	FIBRONECTIN TYPE-III 2.
FT	DISULFID	985	997	BY SIMILARITY.
FT	DISULFID	992	1010	BY SIMILARITY.
FT	DISULFID	1004	1019	BY SIMILARITY.
FT	DISULFID	1024	1038	BY SIMILARITY.
FT	DISULFID	1032	1051	BY SIMILARITY.
FT	DISULFID	1045	1060	BY SIMILARITY.
FT	DISULFID	1065	1077	BY SIMILARITY.
FT	DISULFID	1072	1090	BY SIMILARITY.
FT	DISULFID	1084	1099	BY SIMILARITY.
FT	DISULFID	1105	1117	BY SIMILARITY.
FT	DISULFID	1112	1130	BY SIMILARITY.
FT	DISULFID	1124	1141	BY SIMILARITY.
FT	DISULFID	1145	1155	BY SIMILARITY.
FT	DISULFID	1150	1168	BY SIMILARITY.
FT	DISULFID	1162	1222	BY SIMILARITY.
FT	DISULFID	1232	1244	BY SIMILARITY.
FT	DISULFID	1239	1257	BY SIMILARITY.
FT	DISULFID	1251	1266	BY SIMILARITY.
FT	DISULFID	1275	1288	BY SIMILARITY.
FT	DISULFID	1283	1301	BY SIMILARITY.
FT	DISULFID	1295	1310	BY SIMILARITY.
FT	DISULFID	1333	1351	BY SIMILARITY.
FT	DISULFID	1345	1360	BY SIMILARITY.
FT	CARBOHYD	6	6	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	65	65	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	275	275	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	337	337	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	523	523	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	581	581	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	725	725	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	778	778	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	975	975	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1098	1098	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1152	1152	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1366	1366	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1454	1454	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1514	1514	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	NON_TER	1592	1592	N-LINKED (GLCNAC. .) (POTENTIAL).
SQ	SEQUENCE	1592 AA;	178409 MW;	24EDAA5BA231B203 CRC64;

Query Match 26.1%; Score 52.5; DB 1; Length 1592;
 Best Local Similarity 38.7%; Pred. No. 35;
 Matches 12; Conservative 2; Mismatches 14; Indels 3; Gaps 1;

Qy 1 CSQNEYFDSLHACIPQLRCSSNTPPLTCQ 31
 ||: | : || | | | |
 Db 1378 CSRTEFCQQLHKCIPNWKRCGRR---DCQ 1405

Search completed: June 25, 2002, 16:22:42
 Job time: 627 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 25, 2002, 16:24:25 ; Search time 89.98 Seconds
(without alignments)
65.368 Million cell updates/sec

Title: US-09-854-864-7

Perfect score: 201

Sequence: 1 CSQNEYFDSLHACIPQLRCSSNTPLTCQRYC 34

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_19.*

1: sp_archaea.*

2: sp_bacteria.*

3: sp_fungi.*

4: sp_human.*

5: sp_invertebrate.*

6: sp_mammal.*

7: sp_mhc.*

8: sp_organelle.*

9: sp_phase.*

10: sp_plant.*

11: sp_rodent.*

12: sp_virus.*

13: sp_vertebrate.*

14: sp_unclassified.*

15: sp_rvirus.*

16: sp_bacteriap.*

17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	136	67.7	185	11 088472	088472 mus musculus
2	72.5	36.1	499	11 088714	088714 mus musculus
3	71.5	35.6	249	11 09DBZ3	Q9dbz3 mus musculus
4	71.5	35.6	249	11 09PT35	Q9et35 mus musculus
5	66.5	33.1	293	4 014836	014836 homo sapien
6	64.5	32.1	5374	11 099ND0	Q99nd0 mus musculus
7	61	30.3	966	5 022378	022378 caenorhabdi
8	59.5	29.6	353	5 09VW81	Q9vw81 drosophila
9	59	29.4	937	5 09GTR5	Q9gyr5 caenorhabdi
10	58	28.9	175	11 09SD00	Q9sd00 mus musculus
11	58	28.9	1299	5 026489	Q26489 spodoptera
12	57.5	28.6	62	5 077419	077419 ascaris suu
13	57.5	28.6	1679	5 024301	024301 drosophila
14	57	28.4	321	10 09MAM0	Q9mam0 arabidopsis
15	56	28.4	1717	5 026566	Q26566 schistosoma
16	56	27.9	341	11 09D351	Q9d351 mus musculus

ALIGNMENTS

RESULT 1

088472	PRELIMINARY;	PRT;	185 AA.
ID 088472			
AC 088472			
DT 01-NOV-1998	(TREMBLrel. 08, Created)		
DT 01-NOV-1998	(TREMBLrel. 08, Last sequence update)		
DT 01-JUN-2001	(TREMBLrel. 17, Last annotation update)		
DE B-CELL MATURATION PROTEIN (TUMOR NECROSIS FACTOR RECEPTOR SUPERFAMILY, MEMBER 17).			
DE TNFRSF17.			
GN Mus musculus (Mouse).			
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX NCBI_TaxID=10090;			
RN [1]			
RP SEQUENCE FROM N.A.			
RC STRAIN=BALB/C; TISSUE=SPLEEN;			
RA Madry C., Laabi Y., Callebaut I., Roussel J., Hatzoglou A.,			
RA LeConiat M., Mornon J.-P., Berger R., Tsapis A.;			
RT "Murine BCMA: a new member of the Tumor Necrosis Factor Receptor Superfamily."			
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.			
RN [2]			
RP SEQUENCE FROM N.A.			
RC STRAIN=C57BL/6J; TISSUE=COLON;			
RX MEDLINE=21085660; PubMed=11217851;			
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,			
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,			
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,			
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,			
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,			
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,			
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,			
RA Schriml L.M., Staubli R., Tomita M., Wagner L., Washio T.,			
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,			
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Rinaldo M.F.,			
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,			
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya K., Lee N.H.,			
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,			

Q9pvd4 xenopus lae
Q9ep28 hepatitis c
Q9ve40 drosophila
Q96dn2 homo sapien
Q9zu00 arabidopsis
Q9vit6 drosophila
Q18118 caenorhabdi
Q924k7 meriones un
Q9bus3 homo sapien
Q964d1 entamoeba h
Q964d2 entamoeba h
Q993k9 callitrichi
Q17970 caenorhabdi
Q9ep23 hepatitis c
Q9c091 homo sapien
Q57484 gallus gall
Q75443 homo sapien
Q08523 mus musculu
Q9b107 entamoeba h
Q9nef9 drosophila
Q9w4y4 drosophila
Q9ep09 hepatitis c
Q9tux4 canis famil
Q95326 canis famil
Q95326 canis famil
Q68813 hepatitis c
Q9bps2 bombyx mori
Q9v6r7 homo sapien
Q9f917 comamonas s

17 56 27.9 387 13 Q9PVD4
18 55.5 27.6 146 12 Q9EP28
19 55.5 27.6 225 5 Q9VE40
20 55.5 27.6 955 4 Q96DN2
21 55.5 27.6 989 10 Q9ZU00
22 55.5 27.6 2820 5 Q9VLT6
23 55 27.4 330 5 Q18118
24 55 27.4 344 11 Q924K7
25 55 27.4 450 4 Q9BUS3
26 55 27.4 1074 5 Q964D1
27 55 27.4 1101 5 Q964D2
28 55 27.4 1127 12 Q993K9
29 55 27.4 1513 5 Q17970
30 54.5 27.1 146 12 Q9EP23
31 54.5 27.1 752 4 Q9C091
32 54.5 27.1 1792 13 Q57484
33 54.5 27.1 2155 11 Q08523
34 54.5 27.1 2155 11 Q08523
35 54 26.9 718 5 Q9B107
36 54 26.9 2447 5 Q9NEF9
37 54 26.9 4072 5 Q9W4Y4
38 53.5 26.6 146 12 Q9EP09
39 53.5 26.6 246 6 Q9TUX4
40 53.5 26.6 281 6 Q9Z475
41 53.5 26.6 285 6 Q95326
42 53.5 26.6 414 12 Q68813
43 53.5 26.6 1069 5 Q9BPS2
44 53.5 26.6 2843 4 Q9V6R7
45 53 26.4 305 2 Q9F917

OS	Mus musculus (Mouse).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX	NCBI_TaxID=10090;
[1]	
RN	SEQUENCE FROM N.A.
RP	MDLINE=C57BL/6J; TISSUE=LUNG;
RC	MEDLINE=21085660; PubMed=11217851;
RX	Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA	Kawakita T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA	Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA	Saito T., Okazaki Y., Gotojori T., Bono H., Kasukawa T., Saito R.,
RA	Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA	Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA	Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA	Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA	Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA	Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA	Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA	Gustincic S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA	Lions P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA	Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA	Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA	Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA	Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA	Hayashizaki Y.;
RT	"Functional annotation of a full-length mouse cDNA collection.";
RL	Nature 409:685-690(2001).
DR	EMBL; AK004668; BAB23457.1; -.
DR	MGI; MGI:1889411; Tnfrsf13b.
SQ	SEQUENCE 249 AA; 26981 MW; 6F4290D719FEA037 CRC64;

Query Match	35.6%;	Score 71.5;	DB 11;	Length 249;
Best Local Similarity	35.3%;	Pred. No. 0.0038;		
Matches 12;	Conservative	8;	Mismatches 13;	Indels 1; Gaps 1;

QY	1 CSONEYFDSLHACIPQLCRSSNTPLTGCORYC 34
DB	6 CPKDQYWDSRKSVCVSCALTCQSQR-SQTCTDFC 38
	: :: : : : : : : : :

RESULT 4
Q9ET35 PRELIMINARY; PRT; 249 AA.
ID Q9ET35
AC Q9ET35
DT 01-MAR-2001 (TEMBLrel. 16, Created)
DT 01-MAR-2001 (TEMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
DE TACI PROTEIN.
GN TNFRSF13B.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=21177254; PubMed=10881172;
RX Yan M., Marsters S.A., Grewal I.S., Wang H., Ashkenazi A., Dixit V.M.;
RT "Identification of a receptor for BLyS demonstrates a crucial role in
RT humoral immunity";
RL Nat. Immunol. 1:37-41(2000).
DR EMBL; AF257673; AAG00081.1; -.
DR MGI; MGI:1889411; Tnfrsf13b.
SQ SEQUENCE 249 AA; 26947 MW; CB2F2D61C2931D81 CRC64;

Query Match	35.6%;	Score 71.5;	DB 11;	Length 249;
Best Local Similarity	35.3%;	Pred. No. 0.0038;		
Matches 12;	Conservative	8;	Mismatches 13;	Indels 1; Gaps 1;

QY	1 CSONEYFDSLHACIPQLCRSSNTPLTGCORYC 34
	: :: : : : : : : : :

Db 6 CPKDQYWDSSKSCVSCALTCQSRS-ORTCTDFC 38

```
RESULT 5
O14836 PRELIMINARY; PRT; 293 AA.
AC O14836;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE TRANSMEMBRANE ACTIVATOR AND CAML INTERACTOR.
GN TACI.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97458245; PubMed=93111921;
RA von Bulow G.U., Bram R.J.;
RT "NF-AT activation induced by a CAML-interacting member of the tumor
RL Science 278:138-141(1997).
DR EMBL: AF023614; AAC51790.1; ..
DR InterPro: IPR001368; TNFR_c6.
DR PROSITE: PS00652; TNFR_NGFR_1; UNKNOWN_1.
SQ SEQUENCE 293 AA; 31816 MW; 411799F3DE17A5EB CRC64;

Query Match 33.1%; Score 66.5; DB 4; Length 293;
Best Local Similarity 32.4%; Pred. No. 0.027;
Matches 11; Conservative 8; Mismatches 14; Indels 1; Gaps 1;

QY 1 CSONEYDLSLLHACIPCOLRCSNTPPLTCORYC 34
| : : : | | : : : | : : : | : : : |
Db 34 CPKEQYWDPLLTGCMCKTICNQS-ORTCAAPC 66

RESULT 6
O99ND0 PRELIMINARY; PRT; 5374 AA.
AC O99ND0;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE ZAN (ZONADHESIN).
GN ZAN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=129/SV;
RX MEDLINE=21138439; PubMed=11339002;
RA Wilson M.D., Riemer C., Martindale D.W., Schnupf P., Boright A.P.,
RA Cheung T.L., Hardy D.M., Schwartz S., Scherer S.W., Tsui L.-C.,
RA Miller W., Koop B.F.;
RT "Comparative analysis of the gene-dense ACHE/TFR2 region on human
RT chromosome 7q22 with the orthologous region on mouse chromosome 5.";
RL Nucleic Acids Res. 29:1352-1365(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=129/SV;
RA Cheung T.L., Wilson M.D., Koop B.F., Hardy D.M.;
RT "Genomic Basis of Inter- and Intra-species Variation in Zonadhesin
RT Domain Structure.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 3 MAM DOMAINS.
DR EMBL: AF312033; AAK28924.1; ..
DR EMBL: AY046056; AAL04416.1; ..
DR MGD; MGI:106656; Zan.
DR InterPro: IPR000561; EGF-like.
```

```
DR InterPro: IPR003645; FOLN.
DR InterPro: IPR000998; MAM.
DR InterPro: IPR002919; TIL.
DR InterPro: IPR003328; TILa.
DR InterPro: IPR001846; Vwd.
DR InterPro: IPR001007; VWFC.
DR Pfam; PF00629; MAM; 3.
DR Pfam; PF01826; TIL; 25.
DR Pfam; PF02345; TILa; 25.
DR Pfam; PF00094; vwd; 4.
DR SMART; SM00001; EGF-like; 1.
DR SMART; SM00274; FOLN; 21.
DR SMART; SM00137; MAM; 3.
DR SMART; SM00214; VWC; 25.
DR SMART; SM00216; VWD; 4.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS01186; EGF_2; 18.
DR PROSITE; PS00060; MAM_2; 3.
KW EGF-like domain; Glycoprotein.
SQ SEQUENCE 5374 AA; 579536 MW; 90D2D8CFE5DE24EB CRC64;

Query Match 32.1%; Score 64.5; DB 11; Length 5374;
Best Local Similarity 35.1%; Pred. No. 0.99;
Matches 13; Conservative 5; Mismatches 16; Indels 3; Gaps 2;

QY 1 CSONEYDLSLLHACIP-COLRCSNTP--PLTCORYC 34
| : : : | | : : : | : : : | : : : |
Db 3298 CPTNSQTDCLPSCVPCSNRCEVTSVPSSCREGC 3334

RESULT 7
Q22378 PRELIMINARY; PRT; 966 AA.
AC Q22378;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE HYPOTHETICAL 102.5 KDA PROTEIN.
GN T10E10.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=BRISTOL N2;
RA Gaisel C.;
RT "The sequence of C. elegans cosmid T10E10.";
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=BRISTOL N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; U39644; AAA80360.2; ..
DR HSP; P10969; IWGT.
DR InterPro: IPR002557; Chitin_binding.
DR InterPro: IPR002899; EB.
DR InterPro: IPR000794; Ketoacyl-synt.
DR InterPro: IPR003571; Snake_toxin.
DR Pfam; PF01607; Chitin_bind_2; 2.
DR SMART; SM00289; WRI; 12.
DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; UNKNOWN_1.
```

RA Adams M.D., Celniker S.E., Holl R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Vandell M.D., Zhang Q., Chen L.X.,
 RA Brannon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davidson P.L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fostler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupsky M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RT "The genome sequence of *Drosophila melanogaster*."
 RT Science 287:2185-2195(2000).
 3R EMBL: AE003514; AAF49068.1; -
 DR Flybase: FBgn0036940; CG7348.
 DR InterPro: IPR002557; Chitin_binding.

RESULT	10	
Q9D8D0		
ID	Q9D8D0	PRELIMINARY; PRT; 175 AA.
AC	Q9D8D0;	
DT	01-JUN-2001 (TRENBLRel. 17, Created)	
DT	01-JUN-2001 (TRENBLRel. 17, Last sequence update)	
DT	01-DEC-2001 (TRENBLRel. 19, Last annotation update)	
DE	201006P15RIK PROTEIN (BAFF RECEPTOR).	
OS	Mus musculus (Mouse).	
GN	2010006P15RIK.	
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus	

DR	MEROPS; S08.UPB; -.
DR	InterPro; IPR002174; Furin-like.
DR	InterPro; IPR002099; Peptidase_S8.
DR	InterPro; IPR002884; P_domain.
DR	Pfam; PF01483; P; 1.
DR	Pfam; PF00062; Peptidase_S8; 1.
DR	PRINTS; PR00723; SUBTILISIN.
DR	PRODOM; PD000717; P_domain; 1.
DR	SMART; SM00261; FU; 10.
DR	PROSITE; PS00136; SUBTILASE_ASP; 1.
DR	PROSITE; PS00137; SUBTILASE_HIS; 1.
DR	PROSITE; PS00138; SUBTILASE_SER; 1.
KW	Protease.
SQ	SEQUENCE 1299 AA; 142020 MW; 4C3799C7BBC572AB CRC64;
Query Match 28.9%; Score 58; DB 5; Length 1299;	
Best Local Similarity 37.1%; Pred.No. 2.4;	
Matches 13; Conservative 5; Mismatches 9; Indels	
QY	1 CSQNEYFDSLHACIPCOLRCS-----SNTPLTC 30 :
Db	1150 CSRPLRIDRLNQCVP---CSERGVTNSTPTDC 1181 :
RESULT 12	
O77419	PRELIMINARY; PRT; 62 AA.
ID O77419	
AC O77419	
DT 01-NOV-1998	(TrEMBLrel. 08, Created)
DT 01-NOV-1998	(TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2001	(TrEMBLrel. 17, Last annotation update)
DE CHMOTRYPSIN/ELASTASE INHIBITOR-1 (FRAGMENT).	
GN ASC/E-1.	
OS Ascaris suum (Pig roundworm) (Ascaris lumbricoides);	
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Asca	
OC Ascarididae; Ascaris.	
OX NCBI_TaxID=6253;	
RN [1]	
RP SEQUENCE FROM N.A.	
RX MEDLINE=9829737; PubMed=9635450;	
RA Lu C.C., Nguyen T., Morris S., Hill D., Sakanari J.A.;	
RT "Aniakas simplex: mutational bursts in the reactive site c	
RT serine protease inhibitors from an ascarid nematode.";	
RL Exp. Parasitol. 89:257-261(1998).	
DR EMBL; U94499; AAC61300.1; -.	
DR HSSP; P07851; IEAI.	
DR InterPro; IPR002919; TIL.	
DR Pfam; PF01826; TIL; 1.	
FT NON_TER 1	
FT NON_TER 62	
FT NON_TER 62	
SQ SEQUENCE 62 AA; 6794 MW; 154CE25A375E0B2C CRC64;	
Query Match 28.6%; Score 57.5; DB 5; Length 62;	
Best Local Similarity 37.1%; Pred.No. 0.14;	
Matches 13; Conservative 5; Mismatches 10; Indels	
QY	1 CSQNEYFDSLHACIPCOLRCS--SNTPLTCQR 32 :
Db	5 CGPNEVWTE---CTGCKMKCGDPENTCPFLMCR 35 :
RESULT 13	
ID Q24301	PRELIMINARY; PRT; 1679 AA.
AC Q24301	
DT 01-NOV-1996	(TrEMBLrel. 01, Created)
DT 01-NOV-1996	(TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001	(TrEMBLrel. 19, Last annotation update)
DE FUR2 PROTEIN.	
OS FROSOPHILA melanogaster (Fruit fly).	
GN FUR2 OR DFUR2 OR CG4235 OR CG18734.	

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE EPIDERMAL GROWTH FACTOR RECEPTOR PRECURSOR.
GN SER.
OS Schistosoma mansoni (Blood fluke).
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;
OC Schistosomatoidea; Schistosomatidae; Schistosoma.
OX NCBI_taxid=6183;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PUERTO RICAN;
RX MEDLINE=92365727; PubMed=1501637;
RA Shoemaker C.B., Ramachandran H., Landa A., dos Reis M.G., Stein L.;
RT "Alternative splicing of the Schistosoma mansoni gene encoding a
RL homologue of epidermal growth factor receptor.";
RL Mol. Biochem. Parasitol. 53:17-32(1992).
DR EMBL; M86396; AAA29866.1; -
DR HSSP; P11362; 1FGK
DR InterPro; IPR000494; EGFR_L.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002174; Furin-like.
DR InterPro; IPR001245; Tyr_pkinase.
DR InterPro; IPR000822; Znf-C2H2.
DR Pfam; PF00757; Furin-like; 1.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF01030; Recep_L_domain; 1.
DR SMART; SM00261; FU; 6.
DR SMART; SM00219; TyrKC; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
KW ATP-binding; Receptor; Signal; Transferase; Tyrosine-protein kinase.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 1717 EPIDERMAL GROWTH FACTOR RECEPTOR.
SQ SEQUENCE 1717 AA; 192304 MW; 1101A338C1653D8C CRC64;

Query Match 28.4%; Score 57; DB 5; Length 1717;
Best Local Similarity 40.9%; Pred. No. 4.6;
Matches 9; Conservative 2; Mismatches 11; Indels 0; Gaps 0;

QY 1 CSONEYFDSLHACIPQLRCS 22
Db 646 CPRNTYIDPQTRHCLPCNCS 667
|:|:|:|:|:|:|:|:|

Search completed: June 25, 2002, 16:24:27
Job time: 657 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 25, 2002, 16:16:08 ; Search time 88.08 Seconds
(without alignments)
42.876 Million cell updates/sec

Title: US-09-854-864-7

Perfect score: 201

Sequence: 1 CSONEYFDSLHACIPQLRCSSNTPPLTCQRYC 34

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 11107396 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_032802.*

- 1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	201	100.0	184	21	AA08843	Amino acid sequenc
2	201	100.0	184	21	AA094001	A human BCMA protei
3	201	100.0	184	22	AAE09241	Human BCMA protein
4	201	100.0	184	22	AAE00506	Human B cell matur
5	201	100.0	184	22	AAE06098	Human BAF receptor
6	201	100.0	184	22	AAE00507	Human B cell matur
7	201	100.0	302	22	AAE00507	Human BCMA-Immunog
8	201	100.0	302	22	AAE06099	Mouse IgG signal/h
9	159.5	79.4	157	22	AAE06070	Human BAF receptor
10	136	67.7	185	21	AA08844	Amino acid sequenc
11	136	67.7	185	22	AA071980	Murine B cell matu

12	71.5	35.6	249	21	AA094006	A murine ztnf4, a
13	66.5	33.1	166	19	AAW75785	Human lymphocyte s
14	66.5	33.1	265	22	AAE09244	Human TAC1 splice
15	66.5	33.1	293	19	AAW75783	Human lymphocyte s
16	66.5	33.1	293	21	AA036312	Human neutrophilic a
17	66.5	33.1	293	21	AA094000	A transmembrane ac
18	66.5	33.1	293	22	AAE09240	Human TAC1 protein
19	66.5	33.1	293	22	AAE071914	Human tumour necro
20	59.5	29.6	353	22	ABB71555	Drosophila melanog
21	59.5	29.6	897	22	ABG06309	Novel human diagno
22	59.5	29.6	897	22	ABG65887	Amino acid sequenc
23	59.5	29.6	993	22	AAE06309	Human acid sequenc
24	59.5	29.6	247	21	AAE09398	Human BR43x2, an i
25	57.5	28.6	1679	22	ABB60498	Drosophila melanog
26	57.5	28.6	1679	22	ABB60502	Drosophila melanog
27	56.7	27.9	2476	20	AAW67738	Pig p105 zona pell
28	55.5	27.6	225	22	ABB71511	Drosophila melanog
29	55.5	27.6	2820	22	ABB63296	Drosophila melanog
30	55.5	27.4	955	22	ABG22836	Novel human diagno
31	54.5	26.9	98	21	AAE065429	Human 5' EST relat
32	54.5	26.9	121	22	AAE095477	Human reproductive
33	54.5	26.9	182	22	AAE093696	Human protein sequ
34	54.5	26.9	986	22	ABG07760	Novel human diagno
35	54.5	26.9	4072	22	ABB63614	Drosophila melanog
36	53.5	26.6	508	22	ABG22213	Novel human diagno
37	53.5	26.6	2957	22	ABG22213	Novel human diagno
38	53.5	26.6	5405	16	AAW14749	IgG-Fc binding pro
39	53.5	26.6	7337	22	ABG22216	Novel human diagno
40	53.5	26.4	126	20	AAW98909	Mouse IMC carcino
41	53.5	26.4	131	20	AAW98908	Mouse IMC carcino
42	52.5	26.1	166	22	AAO13678	Human polypeptide
43	52.5	26.1	976	22	ABB11195	Human transmembran
44	52.5	25.9	330	22	ABB62442	Drosophila melanog
45	52.5	25.9	780	22	ABG07994	Novel human diagno

ALIGNMENTS

RESULT 1
AA08843
ID AA08843 standard; peptide; 184 AA.

XX AA08843;
DT 02-JAN-2001 (first entry)
XX
DE Amino acid sequence of human.

XX BCMA; necrosis factor-kB activator; NF-kB; gene expression; cancer;
KW anti-cell death gene; apoptosis; viral infection; inflammatory response;
KW rheumatoid arthritis; inflammatory bowel disease; septic shock.

OS Homo sapiens.

XX Key Location/Qualifiers
FT Domain 57..77
FT /note= "putative transmembrane domain"

XX WO2000050633-A1.

XX 31-AUG-2000.

XX 24-FEB-2000; 2000WO-US04925.

XX 24-FEB-1999; 99US-0121485.

XX (GEO) GEN HOSPITAL CORP.

XX Seed B, Ting A;

XX WPI; 2000-558405/51.

XX

DR WPI: 2001-541628/60.
 DR N-PSDB; AAD15902.
 XX Inhibiting or neutralizing TALL-1 or APRIL polypeptide biological
 PT activity, for treating autoimmune disorders and cancer, comprises
 PT exposing the cells to TALL-1 or APRIL polypeptide agonists or
 PT antagonists -
 XX
 PS Example 2; Fig 2; 160pp; English.
 CC The invention relates to methods of using one or more agonists or
 CC antagonists to modulate the activity of the members of TNF (tumour
 CC necrosis factor) especially TALL-1, APRIL and TNF receptor (TNFR)
 CC e.g. TACI or BCMA. The method is useful for treating pathological
 CC conditions or diseases associated with increased TALL-1 and APRIL
 CC expression or activity. TALL-1 and APRIL antagonists are used to
 CC block the interaction between APRIL and TALL-1 with TACI or BCMA.
 CC They are useful for treating a mammal suffering from cancer such
 CC as leukaemia, lymphoma, myeloma, cancers of lung and colon and
 CC autoimmune diseases e.g. rheumatoid arthritis, multiple sclerosis,
 CC psoriasis and lupus erythematosus. The present sequence is human
 CC BCMA protein.
 XX
 SQ Sequence 184 AA;
 Query Match 100.0%; Score 201; DB 22; Length 184;
 Best Local Similarity 100.0%; Pred. No. 3.4e-18;
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CSONEYFDSLHACIPCQLRCSNTPTPLTCQRYC 34
 DB 8 csqneyfdslhacipcqlrcsntptpltcqryc 41
 RESULT 4
 AAE00506
 ID AAE00506 standard; Protein; 184 AA.
 XX
 AC AAE00506;
 XX
 DT 31-JUL-2001 (first entry)
 XX
 DE Human B cell maturation protein (BCMA).
 KW Human; A Proliferation Inducing Ligand Receptor; APRIL-R; cytostatic;
 KW gene therapy; cancer; nephrotropic; renal disorder; autoimmune disease;
 KW carcinoma; lung; colon; breast; prostate; Grave's disease; hypertension;
 KW systemic lupus erythematosus; SLE; inflammation; cardiovascular disease;
 KW B-cell lympho-proliferative disorder; BCM; immunosuppressive disease;
 KW organ transplantation; HIV; human immunodeficiency virus; TNF;
 KW tumour necrosis factor; BCMA; B cell maturation protein.
 XX
 OS Homo sapiens.
 XX
 XX WO200124811-A1.
 PN
 XX
 XX 12-APR-2001.
 XX
 XX 05-OCT-2000; 2000WO-US27579.
 XX
 XX 06-OCT-1999; 99US-0157933.
 PR 11-FEB-2000; 2000US-0181807.
 PR 30-JUN-2000; 2000US-0215688.
 XX
 XX (BIOJ) BIOGEN INC.
 PA (APOT-) APOTECH R & D SA.
 XX
 XX Schneider P, Thompson J, Cachero T, Ambrose C, Rennert P;
 PI WPI: 2001-266242/27.
 DR N-PSDB; AAD03844.
 DR
 XX

PT Treating a mammal for a condition associated with undesired cell
 PT proliferation such as cancer or carcinoma, comprises administering a
 PT composition comprising A Proliferation Inducing Ligand Receptor
 PT (APRIL-R) antagonist -
 XX
 PS Claim 3; Fig 3A; 85pp; English.
 XX
 CC The invention relates to a method of treating a mammal for a condition
 CC associated with undesired cell proliferation such as cancer or
 CC carcinoma. The method involves administering a composition comprising
 CC A Proliferation Inducing Ligand Receptor (APRIL-R) also referred as
 CC B cell maturation protein (BCM or BCMA) antagonist that antagonises the
 CC interaction between APRIL and its cognate receptor(s). This method is
 CC useful for treating undesired cell proliferation such as cancer or
 CC carcinoma e.g. human lung carcinoma, colon carcinoma, breast carcinoma,
 CC prostate carcinoma, and other carcinomas whose proliferation is modulated
 CC by APRIL. It is also useful for treating autoimmune diseases (Grave's
 CC disease, systemic lupus erythematosus-SLE); hypertension, cardiovascular
 CC diseases, renal disorders, B-cell lympho-proliferative disorders,
 CC immunosuppressive diseases, organ transplantation, inflammation and
 CC human immunodeficiency virus (HIV), and for treating, suppressing or
 CC altering an immune response involving a signalling pathway between
 CC APRIL-R and its ligand. APRIL-R DNA is also useful in gene therapy.
 CC The present sequence is human APRIL-R also referred as BCMA or
 CC BCM protein.
 XX
 SQ Sequence 184 AA;
 Query Match 100.0%; Score 201; DB 22; Length 184;
 Best Local Similarity 100.0%; Pred. No. 3.4e-18;
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CSONEYFDSLHACIPCQLRCSNTPTPLTCQRYC 34
 DB 8 csqneyfdslhacipcqlrcsntptpltcqryc 41
 RESULT 5
 AAB60698
 ID AAB60698 standard; Protein; 184 AA.
 XX
 AC AAB60698;
 XX
 DT 22-MAY-2001 (first entry)
 XX
 DE Human BAFF receptor (BAFF-R).
 KW Human BAFF-R; BAFF receptor; TNF family; immunoregulatory agent;
 KW immune-related disorder; B-cell growth inhibitor; BCMA;
 KW B-cell maturation inhibitor; immunoglobulin production inhibitor;
 KW autoimmune disorder; B-cell lymphoproliferative disorder; hypertension;
 KW renal disorder; immunosuppressive disorder; HIV infection;
 KW organ transplantation; antiinflammatory; systemic lupus erythematosus;
 KW autoimmune haemolytic anaemia; Grave's disease; multiple myeloma;
 KW B-cell carcinoma; leukaemia; rapidly progressive glomerulonephritis;
 KW lymphoma; gene therapy; cancer; tumour.
 XX
 OS Homo sapiens.
 XX
 XX WO200112812-A2.
 PN
 XX
 XX 22-FEB-2001.
 XX
 XX 16-AUG-2000; 2000WO-US22507.
 XX
 XX 17-AUG-1999; 99US-0149378.
 PR 11-FEB-2000; 2000US-0181684.
 PR 18-FEB-2000; 2000US-0183536.
 XX
 XX (BIOJ) BIOGEN INC.
 PA (APOT-) APOTECH R & D SA.
 XX

PI Mackay F, Browning J, Ambrose C, Tschopp J, Schneider P;
 PI Thompson J;
 XX WPI; 2001-202866/20.
 DR N-PSDB; AAF59998.
 XX
 PT Inhibiting dendritic cell-induced B-cell growth, maturation and B-cell
 PT lympho-proliferative disorder by administering BAFF-receptor
 PT polypeptide, chimeric molecule comprising receptor or anti-BAFF-R
 PT antibody homolog
 XX
 XX
 PS Claim 20; Fig 1; 59pp; English.
 XX
 CC The invention relates to the use of a BAFF receptor (BAFF-R, also known
 CC as BCMA) protein, or a BAFF-R fusion protein as an agent for the
 CC treatment of a variety of immune-related disorders. BAFF-R is a member of
 CC the TNF (tumour necrosis factor) family, acting as an immunoregulatory
 CC agent, and also plays a role in the development of hypertension and
 CC related disorders. BAFF-R, fusion proteins containing it, and BAFF-R-
 CC specific antibodies can be used for inhibiting B-cell growth, dendritic
 CC cell-induced B-cell growth and maturation, and immunoglobulin production,
 CC and in the treatment of autoimmune disorders, B-cell lymphoproliferative
 CC disorders, hypertension and renal disorders. The BAFF-R proteins may also
 CC be used in the treatment of immunosuppressive disorders and HIV
 CC infection, and in patients undergoing organ transplantation. The BAFF-R
 CC proteins or BAFF-R specific antibodies may be used for treating,
 CC suppressing or altering an immune response involving a signalling pathway
 CC between BAFF-R and BAFF, thereby inhibiting inflammation. Since BAFF-R
 CC inhibits B-cell growth and maturation it is useful for treating diseases
 CC such as systemic lupus erythematosus, autoimmune haemolytic anaemia,
 CC Grave's disease, multiple myeloma, B-cell carcinomas, leukaemia, rapidly
 CC progressive glomerulonephritis, and lymphomas. Nucleic acids encoding
 CC human BAFF-R may be used in gene therapy to treat tumours, lymphomas,
 CC autoimmune disorders and inherited B-cell-associated disorders. The
 CC present sequence represents human BAFF-R.
 XX
 XX Sequence 184 AA;
 SQ
 Query Match 100.0%; Score 201; DB 22; Length 184;
 Best Local Similarity 100.0%; Pred. No. 3.4e-18;
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CSQNEYFDSLHACIPQLRCSSNTPPLTCQRYC 34
 ||||||||||||||||||||||||||||||||
 Db 8 csqneyfdsllhacipqlrcssntppltcqryc 41
 RESULT 6
 AAY71979
 ID AAY71979 standard; Protein; 184 AA.
 XX
 AC AAY71979;
 XX
 XX 28-MAR-2001 (first entry)
 DT
 XX
 DE Human B cell maturation factor (BCMA) protein.
 XX
 KW Human; Tumour Necrosis Factor; TNF; immunosuppressant; TALL-1;
 KW Tumour necrosis factor and Apol-related Leucocyte expressed Ligand 1;
 KW therapy; autoimmune disorder; rheumatoid arthritis; multiple sclerosis;
 KW systemic lupus erythematosus; SLE; insulin dependent diabetes mellitus;
 KW thrombocytopenia purpura; acute rheumatic fever; Goodpasture's syndrome;
 KW haemolytic anaemia; Grave's disease; myasthenia gravis; chromosome 16;
 KW post-streptococcal glomerulonephritis; polyarteritis nodosa; BCMA;
 KW B cell maturation factor; pemphigus vulgaris; B-lymphocyte proliferation.
 XX
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH 1..62
 FT Domain
 FT /label= Extracellular_domain
 XX

PN WO200068378-A1.
 XX 16-NOV-2000.
 XX
 PF 05-MAY-2000; 2000WO-US12266.
 XX
 XX 06-MAY-1999; 99US-0132892.
 PR 01-MAY-2000; 2000US-0201012.
 PR
 XX (NAJE-) NAT JEWISH MEDICAL & RES CENT.
 XX
 XX Shu HS;
 PI
 XX WPI; 2001-016094/02.
 DR N-PSDB; AAD02125.
 XX
 PT Isolated TALL-1 protein is used to identify compounds that regulate B
 PT lymphocyte proliferation, used to treat B lymphocyte associated
 PT autoimmune disorders
 XX
 XX Claim 37; Page 104-105; 112pp; English.
 PS
 XX The present invention relates to Tumour necrosis factor (TNF) and
 CC Apol-related Leucocyte expressed Ligand 1 (TALL-1) nucleic acid
 CC molecules, proteins (including homologues), and their antibodies. The
 CC interaction in particular relates to methods for regulating the
 CC interaction between TALL-1 and TALL-1 receptors (BCMA referred as B cell
 CC maturation factor) to regulate monocyte, macrophage and B lymphocyte
 CC mediated immune responses. TALL-1 protein is useful for identifying
 CC compounds that regulate B lymphocyte proliferation. It is also useful for
 CC treating B lymphocyte associated autoimmune disorders like rheumatoid
 CC arthritis, systemic lupus erythematosus (SLE), insulin dependent diabetes
 CC mellitus, multiple sclerosis, myasthenia gravis, Grave's disease,
 CC autoimmune haemolytic anaemia, autoimmune thrombocytopenia purpura,
 CC Goodpasture's syndrome, pemphigus vulgaris, acute rheumatic fever,
 CC post-streptococcal glomerulonephritis, or polyarteritis nodosa.
 CC The TALL-1 protein and its corresponding nucleic acid sequence are also
 CC useful in diagnostic assays.
 CC The present sequence is a human B cell maturation factor (BCMA)
 CC protein. It is the receptor for TALL-1 protein. BCMA gene is
 CC located on chromosome 16. In human tissues, BCMA is expressed by
 CC spleen and lymph nodes but not by brain, muscle, heart, lung, kidney,
 CC pancreas, testis and placenta. BCMA mRNA is absent in the pro-B
 CC lymphocyte stage but its expression increases with B lymphocyte
 CC maturation.
 XX
 XX Sequence 184 AA;
 SQ
 Query Match 100.0%; Score 201; DB 22; Length 184;
 Best Local Similarity 100.0%; Pred. No. 3.4e-18;
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CSQNEYFDSLHACIPQLRCSSNTPPLTCQRYC 34
 ||||||||||||||||||||||||||||||||
 Db 8 csqneyfdsllhacipqlrcssntppltcqryc 41
 RESULT 7
 AAE00507
 ID AAE00507 standard; Protein; 302 AA.
 XX
 AC AAE00507;
 XX
 XX 31-JUL-2001 (first entry)
 DT
 XX
 XX Human BCMA-Immunoglobulin G Fc region fusion construct.
 DE
 XX
 KW Human; A Proliferation Inducing Ligand Receptor; APRIL-R; cytostatic;
 KW gene therapy; cancer; nephrotropic; renal disorder; autoimmune disease;
 KW carcinoma; lung; colon; breast; prostate; Grave's disease; hypertension;
 KW systemic lupus erythematosus; SLE; inflammation; cardiovascular disease;
 KW B-cell lympho-proliferative disorder; BCM; immunosuppressive disease;

KW organ transplantation; HIV; human immunodeficiency virus; TNF; murine;
 KW tumour necrosis factor; B cell maturation protein; BCMA; fusion protein;
 KW immunoglobulin G; IgG; Fc region.

XX
 OS Chimeric - Homo sapiens.
 OS Chimeric - Mus sp.

XX Key Location/Qualifiers
 FH Protein 1..22
 FT /label= Signal_peptide
 FT /note= "Derived from murine Ig kappa sequence"
 FT Protein 23..302
 FT /label= Mature_human_BCMA_IgG_Fc_fusion_protein
 FT Region 23..75
 FT /note= "Derived from human BCMA protein"
 FT Region 76..302
 FT /note= "Derived from human IgG Fc region"
 FT Domain 24..302
 FT /label= Cysteine_rich_domain
 FT /note= "Derived from human BCMA"

XX WO200124811-A1.

XX 12-APR-2001.

XX 05-OCT-2000; 2000WO-US27579.

XX 06-OCT-1999; 99US-0157933.

PR 11-FEB-2000; 2000US-0181807.

PR 30-JUN-2000; 2000US-0215688.

XX (BIOJ) BIOGEN INC.

PA (APOT-) APOTECH R & D SA.

XX Schneider P, Thompson J, Cachero T, Ambrose C, Rennert P;

XX WPI; 2001-266242/27.

DR N-PSDB; AAD03847.

XX Treating a mammal for a condition associated with undesired cell
 PT proliferation such as cancer or carcinoma, comprises administering a
 PT composition comprising A Proliferation Inducing Ligand Receptor
 PT (APRIL-R) antagonist

XX Example 1; Fig 3B; 85pp; English.

XX The invention relates to a method of treating a mammal for a condition
 CC associated with undesired cell proliferation such as cancer or
 CC carcinoma. The method involves administering a composition comprising
 CC A Proliferation Inducing Ligand Receptor (APRIL-R) also referred as
 CC B cell maturation protein (BCN or BCMA) antagonist that antagonises the
 CC interaction between APRIL and its cognate receptor(s). This method is
 CC useful for treating undesired cell proliferation such as cancer or
 CC carcinoma e.g. human lung carcinoma, colon carcinoma, breast carcinoma,
 CC prostate carcinoma, and other carcinomas whose proliferation is modulated
 CC by APRIL. It is also useful for treating autoimmune diseases (Grave's
 CC disease, systemic lupus erythematosus-SLE); hypertension, cardiovascular
 CC diseases, renal disorders, B-cell lympho-proliferative disorders,
 CC immunosuppressive diseases, organ transplantation, inflammation and
 CC human immunodeficiency virus (HIV), and for treating, suppressing or
 CC altering an immune response involving a signalling pathway between
 CC APRIL-R and its ligand. APRIL-R DNA is also useful in gene therapy.
 CC The present sequence is a fusion construct containing human APRIL-R
 CC also referred as BCMA or BCM protein, Fc region of human immunoglobulin
 CC G (IgG) and a signal sequence from murine Ig kappa cDNA.

XX Sequence 302 AA;

Query Match 100.0%; Score 201; DB 22; Length 302;
 Best Local Similarity 100.0%; Pred. No. 5,6e-18;
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSQNEYFDSLHACIPQLRCSSTNPPLTCQRYC 34
 Db 31 csqneyfdslhacipqlrccsstnppltcqryc 64

RESULT 8

AAB60699

ID AAB60699 standard; Protein; 302 AA.

XX AAB60699;

XX 22-MAY-2001 (first entry)

XX Mouse IgG signal/human BAFF-R/human IgG Fc fusion protein, BAFF-R-Fc.

XX Human BAFF-R; BAFF receptor; TNF family; immunoregulatory agent;

XX immune-related disorder; B-cell growth inhibitor;

XX B-cell maturation inhibitor; immunoglobulin production inhibitor;

XX autoimmune disorder; B-cell lymphoproliferative disorder; hypertension;

XX renal disorder; immunosuppressive disorder; HIV infection;

XX organ transplantation; antiinflammatory; systemic lupus erythematosus;

XX autoimmune haemolytic anaemia; Grave's disease; multiple myeloma;

XX B-cell carcinoma; leukaemia; rapidly progressive glomerulonephritis;

XX lymphoma; gene therapy; cancer; tumour; IgG Fc; fusion construct.

XX Chimeric - Homo sapiens.

OS Chimeric - Mus sp.

XX WO200112812-A2.

XX 22-FEB-2001.

XX 16-AUG-2000; 2000WO-US22507.

XX 17-AUG-1999; 99US-0149378.

PR 11-FEB-2000; 2000US-0181864.

PR 18-FEB-2000; 2000US-0183536.

XX (BIOJ) BIOGEN INC.

PA (APOT-) APOTECH R & D SA.

XX MacKay F, Browning J, Ambrose C, Tschopp J, Schneider P;

PI Thompson J;

XX WPI; 2001-202866/20.

DR N-PSDB; AAF59999.

XX Inhibiting dendritic cell-induced B-cell growth, maturation and B-cell

PT lympho-proliferative disorder by administering BAFF-receptor

PT polypeptide, chimeric molecule comprising receptor or anti-BAFF-R

PT antibody homolog

XX Example 4; Fig 2; 59pp; English.

XX The invention relates to the use of a BAFF receptor (BAFF-R, also known
 CC as BCMA) protein, or a BAFF-R fusion protein as an agent for the
 CC treatment of a variety of immune-related disorders. BAFF-R is a member of
 CC the TNF (tumour necrosis factor) family, acting as an immunoregulatory
 CC agent, and also plays a role in the development of hypertension and
 CC related disorders. BAFF-R, fusion proteins containing it, and BAFF-R-
 CC specific antibodies can be used for inhibiting B-cell growth, dendritic
 CC cell-induced B-cell growth and maturation, and immunoglobulin production,
 CC and in the treatment of autoimmune disorders, B-cell lymphoproliferative
 CC disorders, hypertension and renal disorders. The BAFF-R proteins may also
 CC be used in the treatment of immunosuppressive disorders and HIV
 CC infection, and in patients undergoing organ transplantation. The BAFF-R
 CC proteins or BAFF-R specific antibodies may be used for treating,
 CC suppressing or altering an immune response involving a signalling pathway
 CC between BAFF-R and BAFF, thereby inhibiting inflammation. Since BAFF-R
 CC inhibits B-cell growth and maturation it is useful for treating diseases
 CC such as systemic lupus erythematosus, autoimmune haemolytic anaemia,
 CC Grave's disease, multiple myeloma, B-cell carcinomas, leukaemia, rapidly
 CC progressive glomerulonephritis, and lymphomas. Nucleic acids encoding

CC human BAFF-R may be used in gene therapy to treat tumours, lymphomas,
 CC autoimmune disorders and inherited B-cell-associated disorders. The
 CC present sequence represents the BAFF-R fusion protein BAFF-R-FC,
 CC comprising a mouse IgG-kappa signal sequence, residues 1-153
 CC of human BAFF-R and a human IgG FC sequence.
 XX
 SQ Sequence 302 AA;

Query Match 100.0%; Score 201; DB 22; Length 302;
 Best Local Similarity 100.0%; Pred. No. 5.6e-18;
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CSQNEYFDSLHACIPQLRCSSNTPPLTCQRYC 34
 |||||
 Db 31 csqneyfdsllhacipqlrcssntppltcqrvc 64

RESULT 9
 AAB60700
 ID AAB60700 standard; Protein: 157 AA.

XX AC AAB60700;

XX DT 22-MAY-2001 (first entry)

XX DE Human BAFF receptor (BAFF-R) sequence encoded by A plasmid pJST535.

XX KW Human BAFF-R: BAFF receptor; TNF family; immunoregulatory agent;
 KW immune-related disorder; B-cell growth inhibitor; BCMA;
 KW B-cell maturation inhibitor; immunoglobulin production inhibitor;
 KW autoimmune disorder; B-cell lymphoproliferative disorder; hypertension;
 KW renal disorder; immunosuppressive disorder; HIV infection;
 KW organ transplantation; antiinflammatory; systemic lupus erythematosus;
 KW autoimmune haemolytic anaemia; Grave's disease; multiple myeloma;
 KW B-cell carcinoma; leukaemia; rapidly progressive glomerulonephritis;
 KW lymphoma; gene therapy; cancer; tumour; plasmid pJST535.

XX OS Homo sapiens.

XX PN WO200112812-A2.

XX PD 22-FEB-2001.

XX PF 16-AUG-2000; 2000WO-US22507.

XX PR 17-AUG-1999; 99US-0149378.

XX PR 11-FEB-2000; 2000US-0181684.

XX PR 18-FEB-2000; 2000US-0183536.

XX PA (BIOJ) BIOGEN INC.

XX PA (APOT-) APOTEC R & D SA.

XX PI Mackay F, Browning J, Ambrose C, Tschopp J, Schneider P;

XX PI Thompson J;

XX DR WPI; 2001-202866/20.

XX DR N-PSDB; AAF60000.

XX XX

XX PS Example 1; Fig 3; 59pp; English.

XX CC The*invention relates to the use of a BAFF receptor (BAFF-R, also known
 CC as BCMA) protein, or a BAFF-R fusion protein as an agent for the
 CC treatment of a variety of immune-related disorders. BAFF-R is a member of
 CC the TNF (tumour necrosis factor) family, acting as an immunoregulatory
 CC agent, and also plays a role in the development of hypertension and
 CC related disorders. BAFF-R, fusion proteins containing it, and BAFF-R-
 CC specific antibodies can be used for inhibiting B-cell growth, dendritic

CC cell-induced B-cell growth and maturation, and immunoglobulin production,
 CC and in the treatment of autoimmune disorders, B-cell lymphoproliferative
 CC disorders, hypertension and renal disorders. The BAFF-R proteins may also
 CC be used in the treatment of immunosuppressive disorders and HIV
 CC infection, and in patients undergoing organ transplantation. The BAFF-R
 CC proteins or BAFF-R specific antibodies may be used for treating,
 CC suppressing or altering an immune response involving a signalling pathway
 CC between BAFF-R and BAFF, thereby inhibiting inflammation. Since BAFF-R
 CC inhibits B-cell growth and maturation it is useful for treating diseases
 CC such as systemic lupus erythematosus, autoimmune haemolytic anaemia,
 CC Grave's disease, multiple myeloma, B-cell carcinomas, leukaemia, rapidly
 CC progressive glomerulonephritis, and lymphomas. Nucleic acids encoding
 CC human BAFF-R may be used in gene therapy to treat tumours, lymphomas,
 CC autoimmune disorders and inherited B-cell-associated disorders. The
 CC present sequence represents a human BAFF-R protein sequence as encoded
 CC by plasmid pJST535. However, this BAFF-R protein sequence is 27 amino
 CC acids shorter than that given in AAB60698.

XX SQ Sequence 157 AA;

Query Match 79.4%; Score 159.5; DB 22; Length 157;
 Best Local Similarity 90.6%; Pred. No. 6.2e-13;
 Matches 29; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

Qy 3 QNEYFDSLHACIPQLRCSSNTPPLTCQRYC 34
 |||||
 Db 7 qneyfdsllhacipqlr---ntppltcqrvc 35

RESULT 10

AAB08844

ID AAB08844 standard; peptide: 185 AA.

XX AC AAB08844;

XX DT 02-JAN-2001 (first entry)

XX DE Amino acid sequence of murine BCMA polypeptide.

XX KW BCMA; necrosis factor-kB activator; NF-kB; gene expression; cancer;
 KW anti-cell death gene; apoptosis; viral infection; inflammatory response;
 KW rheumatoid arthritis; inflammatory bowel disease; septic shock.

XX OS Mus musculus.

XX PH Key Location/Qualifiers

XX FT Domain 47..72

XX FT /note= "putative transmembrane domain"

XX XX WO200050633-A1.

XX XX 31-AUG-2000.

XX PF 24-FEB-2000; 2000WO-US04925.

XX PR 24-FEB-1999; 99US-0121485.

XX XX (GEHO) GEN HOSPITAL CORP.

XX PI Seed B, Ting A;

XX XX WPI; 2000-558405/51.

XX PT Identifying a modulator of gene expression for drug designing, by
 PT contacting a compound library with a cell expressing an anti-cell death
 PT gene and reporter gene, and determining alteration in reporter gene
 PT expression -

XX PS Claim 32; Fig 7b; 53pp; English.

XX CC The present sequence represents a BCMA (not defined) polypeptide. BCMA
 CC is a necrosis factor (NF)-kB activator. The method of the invention is

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 25, 2002, 16:12:12 ; Search time 35.41 Seconds
(without alignments)
23.453 Million cell updates/sec

Title: US-09-854-864-7
Perfect score: 201
Sequence: 1 CSQNEYFDSLHACIPCOLRCSSNTPLTCQRYC 34

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2.6/ptodata/2/1aa/5A_COMB.pep.*
2: /cgn2.6/ptodata/2/1aa/5B_COMB.pep.*
3: /cgn2.6/ptodata/2/1aa/6A_COMB.pep.*
4: /cgn2.6/ptodata/2/1aa/6B_COMB.pep.*
5: /cgn2.6/ptodata/2/1aa/PCTUS_COMB.pep.*
6: /cgn2.6/ptodata/2/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	66.5	33.1	166	2	US-08-810-572A-6
2	66.5	33.1	166	4	US-09-290-333-6
3	66.5	33.1	293	2	US-08-810-572A-2
4	66.5	33.1	293	4	US-09-290-333-2
5	56	27.9	2476	2	US-08-276-967-2
6	53.5	26.6	381	4	US-09-257-580-2
7	53.5	26.6	5405	4	US-08-718-388-9
8	51.5	25.6	63	1	US-08-117-080-12
9	51.5	25.6	63	1	US-08-471-329-12
10	51.5	25.6	63	2	US-08-915-142-12
11	51.5	25.6	77	2	US-08-465-380-40
12	51.5	25.6	77	2	US-08-465-380-40
13	51.5	25.6	77	2	US-08-480-478-33
14	51.5	25.6	77	2	US-08-486-397-4
15	51.5	25.6	77	2	US-08-486-397-40
16	51.5	25.6	77	2	US-08-486-399-4
17	51.5	25.6	77	2	US-08-486-399-40
18	51.5	25.6	77	2	US-08-461-965-4
19	51.5	25.6	77	2	US-08-461-965-40
20	51.5	25.6	77	2	US-08-326-110A-33
21	51.5	25.6	77	2	US-08-634-641-4
22	51.5	25.6	77	2	US-08-634-641-40
23	51.5	25.6	77	3	US-09-249-471-4
24	51.5	25.6	77	3	US-09-249-471-40
25	51.5	25.6	77	3	US-09-249-472-4
26	51.5	25.6	77	3	US-09-249-472-40
27	51.5	25.6	77	3	US-09-249-451-4

28	51.5	25.6	77	3	US-09-249-451-40	Sequence 40, Appl
29	51.5	25.6	77	3	US-08-809-455-4	Sequence 4, Appl
30	51.5	25.6	77	3	US-08-809-455-40	Sequence 40, Appl
31	51.5	25.6	77	3	US-09-249-461-4	Sequence 4, Appl
32	51.5	25.6	77	3	US-09-249-461-40	Sequence 40, Appl
33	51.5	25.6	77	3	US-09-249-448-4	Sequence 4, Appl
34	51.5	25.6	77	3	US-09-249-448-40	Sequence 40, Appl
35	51.5	25.6	81	2	US-08-465-380-7	Sequence 7, Appl
36	51.5	25.6	81	2	US-08-480-478-36	Sequence 36, Appl
37	51.5	25.6	81	2	US-08-486-397-7	Sequence 7, Appl
38	51.5	25.6	81	2	US-08-486-399-7	Sequence 7, Appl
39	51.5	25.6	81	2	US-08-461-965-7	Sequence 7, Appl
40	51.5	25.6	81	2	US-08-326-110A-36	Sequence 36, Appl
41	51.5	25.6	81	3	US-08-634-641-7	Sequence 7, Appl
42	51.5	25.6	81	3	US-09-249-471-7	Sequence 7, Appl
43	51.5	25.6	81	3	US-09-249-472-7	Sequence 7, Appl
44	51.5	25.6	81	3	US-09-249-451-7	Sequence 7, Appl
45	51.5	25.6	81	3	US-08-809-455-7	Sequence 7, Appl

ALIGNMENTS

RESULT 1
US-08-810-572A-6
; Sequence 6, Application US/08810572A
; Patent No. 5969102
; GENERAL INFORMATION:
; APPLICANT: Bram, Richard J.
; APPLICANT: von Bulow, Gotz
; TITLE OF INVENTION: A LYMPHOCYTE SURFACE RECEPTOR THAT BINDS
; TITLE OF INVENTION: CAML, NUCLEIC ACIDS ENCODING THE SAME AND METHODS OF USE
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th
; STREET: Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/08/810,572A
; FILING DATE: 28-FEB-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 1340-1-007
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 166 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; US-08-810-572A-6

Query Match 33.1%; Score 66.5; DB 2; Length 166;

RESULT 5
US-08-276-967-2
; Sequence 2, Application US/08276967
; Patent No. 5851817
; GENERAL INFORMATION:
; APPLICANT: Hardy, Daniel M.
; APPLICANT: Garbers, David L.
; TITLE OF INVENTION: Species-Specific Egg-Binding Proteins of
; TITLE OF INVENTION: Sperm
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P. O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210-4433
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/276,967
; FILING DATE: Submitted Herewith
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Kitchell, Barbara S.
; REGISTRATION NUMBER: 33,928
; REFERENCE/DOCKET NUMBER: UTSD:418\KIT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 713-787-1400
; TELEFAX: 713-789-2679
; TELEX: 79-0924
; INFORMATION FOR SEQ ID NO: 2:

RESULT 7
US-08-718-388-9
; Sequence 9, Application US/08718388
; Patent No. 6271362
; GENERAL INFORMATION:
; APPLICANT: MORIKAWA, MINORU
; APPLICANT: HARADA, NAOKI
; TITLE OF INVENTION: GENE ENCODING IGG FC REGION-BINDING
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH
; STREET: PO BOX 747
; CITY: FALLS CHURCH
; STATE: VA
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/718,388
; FILING DATE:

```
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: MURPHY JR, GERALD M
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 0230-111
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5405 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-718-388-9

Query Match 26.6%; Score 53.5; DB 4; Length 5405;
Best Local Similarity 34.3%; Pred. No. 4.3e+02;
Matches 12; Conservative 5; Mismatches 13; Indels 5; Gaps 2;

Qy 1 CSONEYFDSLLHACI-PCQLRCSSNTPPLTCQRYC 34
Db 2733 CPQNSHYE----LCADTCSLGSALSAPLQCPDGC 2763

RESULT 8
US-08-117-080-12
; Sequence 12, Application US/08117080
; Patent No. 5482928
; GENERAL INFORMATION:
; APPLICANT: DE BOLLE, MIGUEL
; APPLICANT: BROEKAERT, WILLEM F
; APPLICANT: CAMMUE, BRUNO PA
; APPLICANT: VANDERLEYDEN, JOZEF
; APPLICANT: REES, SARAH B
; TITLE OF INVENTION: BIOCIDAL PROTEINS
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DARBY & CUSHMAN
; STREET: 1100 NEW YORK AVENUE, NW; NINTH FLOOR, EAST
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/117,080
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB92/00423
; FILING DATE: 10-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9105052.6
; FILING DATE: 11-MAR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9105684.6
; FILING DATE: 19-MAR-1991
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861 3000
; TELEFAX: (202) 822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 63 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
```

```
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: FIGURE 7 AMINO ACID SEQUENCE MJ-AMP2
US-08-117-080-12

Query Match 25.6%; Score 51.5; DB 1; Length 63;
Best Local Similarity 38.5%; Pred. No. 9.9;
Matches 10; Conservative 4; Mismatches 11; Indels 1; Gaps 1;

Qy 10 LLHACIPQLRCSSNT-PPLTCQRYC 34
Db 24 MIEACIGNGRCNENVGPPYCCSGFC 49

RESULT 9
US-08-471-329-12
; Sequence 12, Application US/08471329
; Patent No. 5689048
; GENERAL INFORMATION:
; APPLICANT: DE BOLLE, MIGUEL
; APPLICANT: BROEKAERT, WILLEM F
; APPLICANT: CAMMUE, BRUNO PA
; APPLICANT: VANDERLEYDEN, JOZEF
; APPLICANT: REES, SARAH B
; TITLE OF INVENTION: BIOCIDAL PROTEINS
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DARBY & CUSHMAN
; STREET: 1100 NEW YORK AVENUE, NW; NINTH FLOOR, EAST
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/471,329
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/117,080
; FILING DATE: 20-DEC-1993
; APPLICATION NUMBER: PCT/GB92/00423
; FILING DATE: 10-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9105052.6
; FILING DATE: 11-MAR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9105684.6
; FILING DATE: 19-MAR-1991
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861 3000
; TELEFAX: (202) 822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 63 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: FIGURE 7 AMINO ACID SEQUENCE MJ-AMP2
US-08-471-329-12

Query Match 25.6%; Score 51.5; DB 1; Length 63;
```

Best Local Similarity 38.5%; Pred. No. 9.9;
Matches 10; Conservative 4; Mismatches 11; Indels 1; Gaps 1;

QY 10 LLHACIPQLRCSSNT-PLPTCQRYC 34
Db 24 MIEACINGGRCNENGVPPYCCSGFC 49

RESULT 10
US-08-915-142-12
; Sequence 12, Application US/08915142
; Patent No. 5942663
; GENERAL INFORMATION:
; APPLICANT: DE BOLLE, MIGUEL
; APPLICANT: BROEKART, WILLEM F
; APPLICANT: CAMMUE, BRUNO PA
; APPLICANT: VANDERLEIDEN, JOZEF
; APPLICANT: REES, SARAH B
; TITLE OF INVENTION: BIOCIDAL PROTEINS
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DARBY & CUSHMAN
; STREET: 1100 NEW YORK AVENUE, NW; NINTH FLOOR, EAST
; STREET: TOWER
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/915,142
; FILING DATE: 20-AUG-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/471,329
; FILING DATE: 02-JUN-1995
; APPLICATION NUMBER: US 08/117,080
; FILING DATE: 20-DEC-1993
; APPLICATION NUMBER: PCT/GB92/00423
; FILING DATE: 10-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9105052.6
; FILING DATE: 11-MAR-1991
; APPLICATION DATA:
; APPLICATION NUMBER: GB 9105584.6
; FILING DATE: 19-MAR-1991
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861 3000
; TELEFAX: (202) 822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 63 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: FIGURE 7 AMINO ACID SEQUENCE MJ-AMP2

US-08-915-142-12

Query Match 25.6%; Score 51.5; DB 2; Length 63;
Best Local Similarity 38.5%; Pred. No. 9.9;
Matches 10; Conservative 4; Mismatches 11; Indels 1; Gaps 1;

QY 10 LLHACIPQLRCSSNT-PLPTCQRYC 34
Db 24 MIEACINGGRCNENGVPPYCCSGFC 49

1; Gaps 1;

QY 10 LLHACIPQLRCSSNT-PLPTCQRYC 34
Db 24 MIEACINGGRCNENGVPPYCCSGFC 49

RESULT 11
US-08-465-380-4
; Sequence 4, Application US/08465380
; Patent No. 5863894
; GENERAL INFORMATION:
; APPLICANT: George P. Vlasuk, Patric H. Stanssens,
; APPLICANT: Joris H.L. Mensens, Marc J. Lauwereys,
; APPLICANT: Yves R. Laroche, Laurent S. Jespers,
; APPLICANT: Yannick G.J. Gansemans, Matthew Moyle,
; APPLICANT: Peter W. Bergum
; TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT
; NUMBER OF SEQUENCES: 356
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/465,380
; FILING DATE: June 5, 1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/326,110
; FILING DATE: October 18, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 213/268
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 77 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Ancylostoma caninum
US-08-465-380-4

US-08-465-380-4

Query Match 25.6%; Score 51.5; DB 2; Length 77;
Best Local Similarity 29.3%; Pred. No. 12;
Matches 12; Conservative 6; Mismatches 12; Indels 11; Gaps 3;

QY 1 CSONEYFDSLHAC---IPCQLRCSSNTP-----PLTCQRYC 34
Db 6 CGENWLDD----CGTKPCEAKCNPEEPEDPICRSRG 42

RESULT 12
US-08-465-380-40
; Sequence 40, Application US/08465380
; Patent No. 5863894
; GENERAL INFORMATION:
; APPLICANT: George P. Vlasuk, Patric H. Stanssens,
; APPLICANT: Joris H.L. Mensens, Marc J. Lauwereys,
; APPLICANT: Yves R. Laroche, Laurent S. Jespers,
; APPLICANT: Yannick G.J. Gansemans, Matthew Moyle,

APPLICANT: Peter W. Bergum
TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT
NUMBER OF SEQUENCES: 356
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/465,380
FILING DATE: June 5, 1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/326,110
FILING DATE: October 18, 1994
ATTORNEY/AGENT INFORMATION:
NAME: BIGGS, SUZANNE L.
REGISTRATION NUMBER: 30,158
REFERENCE/DOCKET NUMBER: 213/268
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 40:
SEQUENCE CHARACTERISTICS:
LENGTH: 77 amino acids
TYPE: amino acid
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Ancylostoma caninum
US-08-465-380-40

Query Match 25.6%; Score 51.5; DB 2; Length 77;
Best Local Similarity 29.3%; Pred. No. 12;
Matches 12; Conservative 6; Mismatches 12; Indels 11; Gaps 3;

Qy 1 CSONEYFDSLHAC---IPCQLRCSSNTP----PLTCQRYC 34
| :||: | | | :||: | | : | |
Db 6 CGENEWLDD----CGTQKPCAKCNEEPPEDPICRSRG 42

RESULT 13

US-08-480-478-33
Sequence 33, Application US/08480478
Patent No. 5864009
GENERAL INFORMATION:
APPLICANT: GEORGE P. VLASUK; PATRICK ERIC
APPLICANT: HUGO STANSSENS; TORIS HILDA
APPLICANT: LIEVEN MESSENS; MARC JOZEF
APPLICANT: LAURENTS; YVES RENE LAROCHE;
APPLICANT: LAURENT STEPHANE JESPEERS; and
APPLICANT: GANSEMAN
TITLE OF INVENTION: NEMATODE-EXTRACTED ANTI-
TITLE OF INVENTION: COAGULANT PROTEIN
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Los Angeles

STATE: California
COUNTRY: U.S.A.
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,478
FILING DATE: 06-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/326,110
FILING DATE: 18 OCTOBER 1994
ATTORNEY/AGENT INFORMATION:
NAME: BIGGS, SUZANNE L.
REGISTRATION NUMBER: 30,158
REFERENCE/DOCKET NUMBER: 208/290
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 77 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-480-478-33

Query Match 25.6%; Score 51.5; DB 2; Length 77;
Best Local Similarity 29.3%; Pred. No. 12;
Matches 12; Conservative 6; Mismatches 12; Indels 11; Gaps 3;

Qy 1 CSONEYFDSLHAC---IPCQLRCSSNTP----PLTCQRYC 34
| :||: | | | :||: | | : | |
Db 6 CGENEWLDD----CGTQKPCAKCNEEPPEDPICRSRG 42

RESULT 14

US-08-486-397-4
Sequence 4, Application US/08486397
Patent No. 5866542
GENERAL INFORMATION:
APPLICANT: George P. Vlasuk, Patric H. Stanssens,
APPLICANT: Joris H.L. Menssens, Marc J. Lauwereys,
APPLICANT: Yves R. Laroche, Laurent S. Jespers,
APPLICANT: Yannick G.J. Gansemans, Matthew Moyle,
APPLICANT: Peter W. Bergum
TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT
TITLE OF INVENTION: PROTEIN
NUMBER OF SEQUENCES: 357
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/486,397
FILING DATE: June 5, 1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 25, 2002, 16:13:22 ; Search time 52.94 Seconds
(without alignments)
301.300 Million cell updates/sec

Title: US-09-854-864-15
Perfect score: 909
Sequence: 1 MSGLGRSRGRSRVDQEER.....SPALPGLKLSADQVALVYST 166

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues
Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	97.5	10.7	1548	2 S34583	serine proteinase
2	86.5	9.5	353	2 T06466	cathepsin B-like c
3	86	9.5	1299	2 T43251	furin (EC 3.4.21.7
4	84.5	9.3	355	2 AD0299	conserved hypothet
5	83	9.1	13288	2 T03099	mucin, submaxillar
6	82.5	9.1	899	2 G02428	subtilisin-like pr
7	82.5	9.1	915	2 JC6148	subtilisin-like pr
8	82	9.0	932	2 I52527	PACE4A - mouse (fr
9	82	9.0	1650	2 S53457	dominant autoantig
10	82	9.0	4660	2 T42737	gp330 protein prec
11	81	8.9	773	1 JH0609	protein-tyrosine-p
12	81	8.9	1124	1 I58488	protein-tyrosine k
13	80.5	8.9	937	2 I53282	gene PACE4 protein
14	80	8.8	108	2 H69834	hypothetical prote
15	80	8.8	330	2 T25169	hypothetical prote
16	80	8.8	996	2 JF0237	apolipoprotein E r
17	79.5	8.7	1984	2 T1371	probable vitelloge
18	79	8.7	269	2 C69312	myodopterin oxid
19	78.5	8.6	630	2 A49656	estrogen-responsiv
20	78.5	8.6	1170	2 A53612	laminin B1k chain
21	78	8.6	596	2 A45664	variant-specific s
22	77.5	8.5	230	2 T19364	hypothetical prote
23	77.5	8.5	2195	2 T34264	hypothetical prote
24	77.5	8.5	2871	2 A55624	fibrillin-1 precur
25	77	8.5	620	2 AF0273	probable iron-sulf
26	77	8.5	751	2 T15230	hypothetical prote
27	77	8.5	1184	2 T09484	cartilage intermed
28	76.5	8.4	1680	2 A43434	furin (EC 3.4.21.7
29	76.5	8.4	1872	2 JC4976	plexin 3 precursor

30 76 8.4 367 2 T45812
31 76 8.4 574 2 B88465
32 76 8.4 915 2 B48225
33 76 8.4 1748 2 S42136
34 75.5 8.3 427 2 S38032
35 75.5 8.3 1620 2 T27283
36 75 8.3 962 2 JC5571
37 75 8.3 969 1 A39490
38 75 8.3 975 2 JC5570
39 75 8.3 2664 2 T28626
40 75 8.3 3002 2 A47221
41 75 8.3 4544 1 S02392
42 75 8.3 4545 1 S25111
43 74.5 8.2 382 1 S48748
44 74.5 8.2 501 2 I61512
45 74.5 8.2 773 2 I46059

ALIGNMENTS

RESULT 1

S34583

Serine proteinase (EC 3.4.21.-) PC6B - mouse

C;Species: Mus musculus (house mouse)

C;Date: 02-Dec-1993 #sequence_revision 10-Nov-1995 #text_change 05-Nov-1999

C;Accession: S34583

R;Nakagawa, T.; Murakami, K.; Nakayama, K.

FEBS Lett. 327, 165-171, 1993

A;Title: Identification of an isoform with an extremely

A;Reference number: S34583; MUID:93327934

A;Accession: S34583

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-1548 <NAK>

A;Cross-references: GB:DL7583; NID:g407344; PIDN:BAA04507.1; PID:d1005033; PID:g44037

C;Keywords: hydrolase; serine proteinase

Query Match 10.7% Score 97.5; DB 2; Length 1548;
Best Local Similarity 30.2%; Pred. No. 1;
Matches 29; Conservative 12; Mismatches 32; Indels 23; Gaps 7;

Qy 14 RVDQ--EERFQGLWTGVAMRSPCEQYWDPLLG-TCMSCRTICN-----HQSQRTCAAF 65

Db 627 RADKHGQERF---LYHGCELENCPVGHY--PAKGHTCLPCPDNCELCYNPHICSRMSGY 681

Qy 66 -----CRSLSCRKEGKFYDHLRLDCISCASIC 93

Db 682 VIIPNHTCQKLECR--QGEFQDSEYECMPCEGC 715

RESULT 2

T06466

cathepsin B-like cysteine proteinase (EC 3.4.22.-) (clone A116) - wheat (fragment)

C;Species: Triticum aestivum (common wheat)

C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-May-2000

C;Accession: T06466

R;Cejudo, F. J.; Murphy, G.; Chinoy, C.; Baulcombe, D.C.

Plant J. 2, 937-948, 1992

A;Title: A gibberellin-regulated gene from wheat with sequence homology to cathepsin

A;Reference number: Z15659; MUID:93258430

A;Accession: T06466

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-353 <CEJ>

A;Cross-references: EMBL:X66013; NID:g21698; PIDN:CAA46811.1; PID:g21699

A;Experimental source: cv. Chinese Spring, etiolated shoots

C;Genetics:

A;Note: Intron positions not resolved (incomplete sequence)

C;Superfamily: papain

C;Keywords: cysteine proteinase; glycoprotein; hydrolase


```

Db 13075 EEKSPGDVWTANCHKCTCTEAKTVDCPKCEPSPCTKGERLIKFKANDTC--CEGCH 13131
Qy 71 CRK-----BQKEYDHLRLDCI--SCASI-----CGQH-----PKQ 99
Db 13132 CEKRTCLFNNTDYEVGSSFDDPNNPCVTYSCONTGFTAVQNCPKQTWCAEDRVYDSKQ 13191
Qy 100 CAYFCENKLR-SPVNL 114
Db 13192 CCYTCKSKCKSPVNV 13207

```

```

RESULT      6
G02428      subtilisin-like proprotein convertase (EC 3.4.21.-) 5 precursor - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 17-Mar-2000
C:Accession: G02428
R:Reudelhuber, T. L.
  submitted to the EMBL Data Library, February 1996
A:Reference number: H01242
A:Accession: G02428
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-899 <REU>
A:Cross-references: EMBL:U49114; NID:g1218057; PIDN:AAA91807.1; PID:g1218058
C:Genetics:
A:Gene: PCS
C:Superfamily: subtilisin-like proteinase PACE4; subtilisin homology
C:Keywords: hydrolase; serine proteinase
F:148-386/Domain: subtilisin homology <SPT>

```

```

Query Match      9.1%; Score 82.5; DB 2; Length 899;
Best Local Similarity 26.0%; Pred. No. 14;
Matches 27; Conservative 10; Mismatches 34; Indels 33; Gaps 7;

QY 27 TGVAMRSCP EYWDPLLLGTGCM SKKTCN HOSQRTCAAF-----CR-----SLSC 71
   | : | : | | | | | | | | | | | | | | | | | |
Db 703 TNSCVTHCPDGSYQDTKKNLKRCSENC-----KTCTEFHNCTECRGLSLQGSRCVSVC 757
   | : | : | | | | | | | | | | | | | | | | | |

QY 72 RKEOGKGYDHLRLDLCISASC-----GQHPKQC-----AYFCEN 106
   | : | : | | | | | | | | | | | | | | | | | |
Db 758 --EDGRYFNG--QDCQPCHRFCATCAGAGADGCINCTEGYFME 797
   | : | : | | | | | | | | | | | | | | | | | |

```

RESULT 7
JC6148
subtilisin-like proprotein convertase (EC 3.4.21.-) homolog - human
N:Alternate names: PC6A protease
C:Species: Homo sapiens (man)
C:Date: 11-Apr-1997 #sequence_revision 09-May-1997 #text_change 22-Jun-1999
C:Accession: JC6148
R:Miranda, L.; Wolf, J.; Pichuantes, S.; Duke, R.; Franzusoff, A.
Proc. Natl. Acad. Sci. U.S.A. 93, 7695-7700, 1996
A:A:Title: Isolation of the human PC6 gene encoding the putative host protease for HIV-1
A:Reference number: JC6148; MUID:96353880
A:Contents: CEM T-cell
A:Accession: JC6148
A:Molecule type: mRNA
A:Residues: 1-915 <MIR>
A:C:Cross-references: GB:U56387; NID:g1498312; PIDN:AAC50643.1; PID:g1498313
C:Comment: This protein functions as a soluble enzyme within the secretory pathway. It
C:Genetics:
A:Gene: pc6A
C:Superfamily: subtilisin-like proteinase PACE4; subtilisin homology
C:Keywords: glycoprotein; hydrolase; serine proteinase
F:164-402/Domain: subtilisin homology <SPT>

Query Match 9.1%; Score 82.5; DB 2; Length 915;
Best Local Similarity 26.0%; Pred. No. 14;
Matches 27; Conservative 10; Mismatches 34; Indels 33; Gaps 7;

```

QY 27 TGVMRSCPPEQYWDPLGLGTCMSCKTICNHOSQRTCAAF-----SLSC 71
Db 719 TNSCVTHCPDGSQDTKKNLCRKSENC-----KTCTEFHNCTECRDLSLQSGRCSVSC 773
QY 72 RKEQGFYDHLRLDRCISASC-----GQHPKQC-----AYFCEN 105
Db 774 --EDGRIYNG--QDCQPCHRFCATCAGAGADGCINCTEGYFMD 813
RESULT 8
PACE4A - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 16-Jul-1999
C:Accession: I52527
R:Hosaka, M.; Murakami, K.; Nakayama, K.
Biomed. Res. 15, 383-390, 1994
A:Title: PACE4A is a ubiquitous endoprotease that has similar but not identical sub-
A:Reference number: I52527
A:Accession: I52527
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-932 <RES>
A:Cross-references: GB:D50060; NID:g769700; PIDN:BAA08777.1; PID:g769701
C:Superfamily: subtilisin-like proteinase PACE4; subtilisin homology
F:I72-410/Domain: subtilisin homology <SRT>

Query Match 9.0%; Score 82; DB 2; Length 932;
Best Local Similarity 25.0%; Pred. No. 15;
Matches 28; Conservative 16; Mismatches 44; Indels 24; Gaps 5;

QY 16 DOEERFPQGLITGVAMRSCPPEQYWDPLGLGTCMSCK--TICNHOSQRTCAAFCSLSCK 73
Db 650 DEEE-----YTGVCHEPCGDKGCDGNADCLNCVHFSLGNSKTNKCVSEC----- 696
QY 74 EOGKFYDHLRLDRCISASC-----GQHPKQC-----AYFCENKLRSPVNLPP 116
Db 697 PLGYFGDAARRCRRCHKGCETCTGRSPAOLSCRRGFGYHHOETNTGVTLC 748

```

RESULT	9	
S53457	dominant autoantigen gp 330 - rat (fragment)	
C:Species:	Rattus norvegicus (Norway rat)	
C:Date:	01-Aug-1995	sequence_revision 01-Sep-1995 #text_change 18-Aug-2000
C:Accession:	S53457	
R:Jokhadze, G.G.;	Oleinikov, A.V.;	Kanalas, J.J.;
Biochem. J.	305, 711-713, 1995	Makker, S.P.
A:Title:	Different molecular forms of rat kidney gp330, the dominant autoantigen of a	
A:Reference number:	S53457; MUID:95151000	
A:Accession:	S53457	
A:Status:	preliminary; nucleic acid sequence not shown	
A:Molecule type:	mRNA	
A:Residues:	1-1650 <JG>	
C:Superfamily:	alpha-2-macroglobulin receptor; JGF homolog; LDL receptor ligand-binding	
F:29-65/Domain:	LDL receptor ligand-binding repeat homology <LDL1>	
F:72-106/Domain:	LDL receptor ligand-binding repeat homology <LDLC>	
F:111-147/Domain:	EGF homology <EG1>	
F:153-188/Domain:	EGF homology <EG2>	
F:196-235/Domain:	LDL receptor YWTD-containing repeat homology <YW01>	
F:236-278/Domain:	LDL receptor YWTD-containing repeat homology <YW02>	
F:279-329/Domain:	LDL receptor YWTD-containing repeat homology <YW03>	
F:330-373/Domain:	LDL receptor YWTD-containing repeat homology <YW04>	
F:374-414/Domain:	LDL receptor YWTD-containing repeat homology <YW05>	
F:415-457/Domain:	LDL receptor YWTD-containing repeat homology <YW06>	
F:466-505/Domain:	EGF homology <EG3>	
F:509-545/Domain:	LDL receptor ligand-binding repeat homology <LDL2>	
F:550-586/Domain:	LDL receptor ligand-binding repeat homology <LDLA>	
F:595-631/Domain:	LDL receptor ligand-binding repeat homology <LDL3>	
F:636-672/Domain:	LDL receptor ligand-binding repeat homology <LDL4>	
F:679-715/Domain:	LDL receptor ligand-binding repeat homology <LDLB>	
F:720-755/Domain:	LDL receptor ligand-binding repeat homology <LDL5>	

F:760-794/Domain: LDL receptor ligand-binding repeat homology <LDL6>
F:799-833/Domain: LDL receptor ligand-binding repeat homology <LDL6>
F:843-879/Domain: LDL receptor ligand-binding repeat homology <LDL7>
F:884-921/Domain: LDL receptor ligand-binding repeat homology <LDLE>
F:929-963/Domain: LDL receptor ligand-binding repeat homology <LDL8>
F:971-1006/Domain: EGF homology <EG4>
F:1012-1048/Domain: EGF homology <EG5>
F:1055-1099/Domain: LDL receptor WYTD-containing repeat homology <YW07>
F:1100-1154/Domain: LDL receptor WYTD-containing repeat homology <YW08>
F:1105-1188/Domain: LDL receptor WYTD-containing repeat homology <YW09>
F:1189-1232/Domain: LDL receptor WYTD-containing repeat homology <YW10>
F:1234-1273/Domain: LDL receptor WYTD-containing repeat homology <YW11>
F:1274-1316/Domain: LDL receptor WYTD-containing repeat homology <YW12>
F:1326-1359/Domain: EGF homology <EG6>

Query Match 9.08; Score 82; DB 2; Length 1650;
Best Local Similarity 25.88; Pred. No. 25;
Matches 40; Conservative 15; Mismatches 64; Indels 36; Gaps 8;

Qy 33 SCPEEQY-WDPLLTGTCMSCKTICNHQ---SORTCAAFCSRSLSCRKEQGFYDH----- 81
Db 71 TCPLHQFRCND--GHCIEMGVRCVNHVDDCSNDEKGCINECLDSSISRCDHNCNTDTIT 128

Qy 82 -----LLRDCISCASI--CGQHPKOCAYFCENKLRSPV-NLPPELRQRSGEVE 127
Db 129 SFYCSCLPGYKLMSDKRSCVDIDECKESPQLCSQKCNVGVSYICKCAPGYIREPDGKSC 188

Qy 128 NNSDN-----SGRY--QGLEHGRGSEASPALPGL 153
Db 189 RQNSNIEPYLIFSRYIRNLTDTGSSYSLLIQL 223

RESULT 10
T42737
gp330 protein precursor - rat
N:Alternate names: megalin
C:Species: Rattus norvegicus (Norway rat)
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 04-Mar-2000
C:Accession: T42737
R:Saito, A.; Pietromonaco, S.; Loo, A.K.C.; Farquhar, M.G.
Proc. Natl. Acad. Sci. U.S.A. 91, 9725-9729, 1994
A:Title: Complete cloning and sequencing of rat gp330/megalin, a distinctive member of the gp330 protein family
F:1-25/Domain: signal sequence #status predicted <SIG>
F:26-4660/Product: gp330 protein #status predicted <MAT>

A:Accession: T42737
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-4660 <SAI>
A:Cross-references: EMBL:L34049; NID:9561852; PID:9561853; PIDN:AAA51369.1
A:Experimental source: strain Sprague-Dawley; kidney
C:Superfamily: alpha-2-macroglobulin receptor; EGF homology; LDL receptor ligand-binding
F:1-25/Domain: signal sequence #status predicted <SIG>
F:26-4660/Product: gp330 protein #status predicted <MAT>

Query Match 9.08; Score 82; DB 2; Length 4660;
Best Local Similarity 25.88; Pred. No. 63;
Matches 40; Conservative 15; Mismatches 64; Indels 36; Gaps 8;

Qy 33 SCPEEQY-WDPLLTGTCMSCKTICNHQ---SORTCAAFCSRSLSCRKEQGFYDH----- 81
Db 3076 TCPLHQFRCND--GHCIEMGVRCVNHVDDCSNDEKGCINECLDSSISRCDHNCNTDTIT 3133

Qy 82 -----LLRDCISCASI--CGQHPKOCAYFCENKLRSPV-NLPPELRQRSGEVE 127
Db 3134 SFYCSCLPGYKLMSDKRSCVDIDECKESPQLCSQKCNVGVSYICKCAPGYIREPDGKSC 3193

Qy 128 NNSDN-----SGRY--QGLEHGRGSEASPALPGL 153
Db 3194 RQNSNIEPYLIFSRYIRNLTDTGSSYSLLIQL 3228

RESULT 11

JH0609
protein-tyrosine-phosphatase (EC 3.1.3.48) P19 - mouse
N:Alternate names: protein-tyrosine-phosphatase PTPY43
C:Species: Mus musculus (house mouse)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: JH0609; PS0365; PS0369; PS0366; G61180
R:den Hertog, J.; Pals, C.E.G.M.; Jonk, L.J.C.; Kruijjer, W.
Biochem. Biophys. Res. Commun. 184, 1241-1249, 1992
A:Title: Differential expression of a novel murine non-receptor protein tyrosine phosphatase
A:Reference number: JH0609; MUID:92272714
A:Accession: JH0609
A:Molecule type: mRNA
A:Residues: 1-773 <DEN>
A:Cross-references: GB:X63440; GB:S36169; NID:9416181; PIDN:CAA45037.1; PID:9416182
A:Experimental source: embryonic carcinoma cell, P19 cell
A:Accession: PS0365
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 88-91,'G',93-110,'G',112-118,'S',120,'T',122 <DE2>
A:Experimental source: embryonic carcinoma cell, P19 cell, clone PTP33
A:Accession: PS0369
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 88-91,'G',93-109,'LG',112-120,'T',122 <DE3>
A:Experimental source: embryonic carcinoma cell, P19 cell, clone PTP59
A:Accession: PS0366
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 88-91,'KY',94-109,'LA',112-118,'S',120-122 <DE4>
A:Experimental source: embryonic carcinoma cell, P19 cell, clone PTP42
R:Yi, T.; Cleveland, J.L.; Ihle, J.N.
Blood 78, 2222-2228, 1991
A:Title: Identification of novel protein tyrosine phosphatases of hematopoietic cells
A:Reference number: A61180; MUID:92032882
A:Accession: G61180
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 124-127,'I',129-229 <YIA>
C:Comment: This protein is located in the cytoplasm.
C:Superfamily: protein-tyrosine-phosphatase, nonreceptor type 12; protein-tyrosine-phosphatase
C:Keywords: phosphoprotein; phosphoric monoester hydrolase; tyrosine-specific phosphatase
F:53-239/Domain: phosphatase catalytic domain #status predicted <PCD>
F:58-232/Domain: protein-tyrosine-phosphatase homology <PTP>
F:231/Active site: Cys (phosphocysteine intermediate) #status predicted
F:237/Binding site: substrate phosphate (Arg) #status predicted

Query Match 8.98; Score 81; DB 1; Length 773;
Best Local Similarity 25.48; Pred. No. 16;
Matches 45; Conservative 18; Mismatches 48; Indels 66; Gaps 11;

Qy 37 EOYWDPLIG----TCMSCKTICNHQSQRTCAAFCSR--LSCRKQGGKPY-----DH- 81
Db 144 ERYW-PLYGEDPITFAFPKISCENEQART-DYFIRTLLEFQNESRRLLYQPHYVWPDHD 201

Qy 82 -----LLRD-----CISCASICQHPKQOCAY-FCENKLR-----S 110
Db 202 VPSSFDSTLDMISLMRYQEHEDVPICITCSAGCGRTGAICAIDYTNLLKAGRIPEFN 261

Qy 111 PVNLPPELRQRSGEVE-----NNSDNGRYQGLEHGRGSEASPALP 151
Db 262 VFNLQEMRTQRHSVOTKQEQYELVHRAIAQLFENSYNCKMF--MEHRRSVVWVKLP 316

RESULT 12
I58388
protein-tyrosine kinase (EC 2.7.1.112), receptor type tek precursor - human
C:Species: Homo sapiens (man)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 16-Jul-1999
C:Accession: I58388
R:Ziegler, S.F.; Bird, T.A.; Schneringer, J.A.; Schooley, K.A.; Baum, P.R.
Oncogene 8, 663-670, 1993
A:Title: Molecular cloning and characterization of a novel receptor protein tyrosine

A:Reference number: I58388; MUID:93173509
A:Accession: I58388
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1124 <RES>
A:Cross-references: GB:L06139; NID:g292823; PIDN:AAA61139.1; PID:g292824
C:Genetics:
A:Gene: GDB:TEK
A:Cross-references: GDB:344185; OMIM:600221
A:Map position: 9p21-9p21
C:Function:
A:Description: catalyzes the phosphorylation of a peptidyl tyrosine residue by ATP
C:Superfamily: protein-tyrosine kinase, receptor type tie; EGF homology; fibronectin ty
C:Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phospho
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-1124/Product: protein-tyrosine kinase, receptor type tek #status predicted <MAT>
F:37-104/Domain: immunoglobulin homology <IM1>
F:135-137/Region: cell attachment (R-G-D) motif
F:211-251/Domain: EGF homology <EG1>
F:255-298/Domain: EGF homology <EG2>
F:302-340/Domain: EGF homology <EG3>
F:364-426/Domain: immunoglobulin homology <IM2>
F:447-527/Domain: fibronectin type III repeat homology <FN3A>
F:542-625/Domain: fibronectin type III repeat homology <FN3B>
F:638-720/Domain: fibronectin type III repeat homology <FN3C>
F:752-772/Domain: transmembrane #status predicted <TMW>
F:822-1099/Domain: protein kinase homology <KIN>
F:830-838/Region: protein kinase ATP-binding motif
F:140-158,399,438,464,560,596,649,691/Binding site: carbohydrate (Asn) (covalent) #statu
F:855,872,964/Active site: Lys, Glu, Asp #status predicted

Query Match 8.9%; Score 81; DB 1; Length 1124;
Best Local Similarity 24.2%; Pred. No. 22;
Matches 39; Conservative 15; Mismatches 63; Indels 44; Gaps 8;

QY 26 WTGVAMRSCPEEQYWD-LLGTGCM-SCKTTCN-HQSQ 59
Db 203 FTRLIVRRC-EAOKWGPCECHLCTACNNGVCHDGTGECIPPGFMGRTCEKACELHTFG 261
QY 60 RTCAAFCSLSCKRQKQKFDHLLRDCISACICGQHPKQCAFCENKLRSPVNLPELR 119
Db 262 RTCKERCSG---QEGCKSVVFLPDYPYGCSCATGWKGLQCNACH-----PGFYGPDCK 312
QY 120 QRSGEVENSDNSGRYQGLEHGRGSEASPALGKLISADQV 160
Db 313 LRCSG---NNGCMCDRFQ-----GCLCSGQWQGLQGEREGI 345

RESULT 13
I53282
gene PACE4 protein - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 22-Jun-1999
C:Accession: I53282
R:Johnson, R.C.; Darlington, D.N.; Hand, T.A.; Bloomquist, B.T.; Mains, R.E.
Endocrinology 135, 1178-1185, 1994
A:Title: PACE4: a subtilisin-like endoprotease prevalent in the anterior pituitary and
A:Reference number: I53282; MUID:94349873
A:Accession: I53282
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-937 <RES>
A:Cross-references: GB:L31894; NID:g496221; PIDN:AAA61987.1; PID:g496222
C:Genetics:
A:Gene: PACE4
C:Superfamily: subtilisin-like proteinase PACE4; subtilisin homology
F:177-415/Domain: subtilisin homology <SBT>

Query Match 8.9%; Score 80.5; DB 2; Length 937;
Best Local Similarity 24.1%; Pred. No. 21;
Matches 27; Conservative 17; Mismatches 45; Indels 23; Gaps 5;

QY 16 DQERFPQGLWTGVAMRSCPEEQYWDPLLGTCMSCK--TICNHQSQRCTCAAFCSLSCKR 73
Db 654 EEEBEY-----TGVCHEPCGDKGCDGSPADQCLNCVHFLSGLNKRKCVSEC----- 701
QY 74 EQGKFYDHLRLDCISCASIC-----GQHPKQC-----AYFCENKLRSPVNLPP 116
Db 702 PLGYFGDTAARRCRCHKGCTCTGSRPTQCLSCRRGFYHHQETNTVTLCP 753

RESULT 14
H69834
hypothetical protein yhjQ - Bacillus subtilis
C:Species: Bacillus subtilis
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 15-Oct-1999
C:Accession: H69834
R:Kunst, F.; Ogawawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber
C.: Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;
A.: Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari,
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gal
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M
Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardino
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mau
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portete
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scan
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Se
akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpstra, P.; Tononi, A.; Tosato, V.; Uchiya
T.: Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida
A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtil
A:Reference number: A69580; MUID:98044033
A:Accession: H69834
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-108 <KUN>
A:Cross-references: GB:Z99109; GB:AL009126; NID:g2633260; PIDN:CAB12900.1; PID:ell830
A:Experimental source: strain 168
C:Genetics:
A:Gene: yhjQ

Query Match 8.8%; Score 80; DB 2; Length 108;
Best Local Similarity 23.6%; Pred. No. 3.5;
Matches 25; Conservative 9; Mismatches 26; Indels 46; Gaps 5;

QY 37 EOYWDPLLGTCMSCKTTCNH-----QSORTCA----- 63
Db 2 EOYSEACIEACIDOMKACNHCFTKCLEESVQHHLSCGIRLDRECAIDCALAVKAMQTDSP 61
QY 64 -----AFCRSL--SCRKQKGF-YDHLRLDCISCASICGQHPKQC 100
Db 62 FMKEICALCADICEACGTEGCKHDHDH---CQACAKACFTCAEQC 103

RESULT 15
T25169
hypothetical protein T23F1.6 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
C:Accession: T25169
R:Wilkinson, J.
submitted to the EMBL Data Library, October 1996
A:Reference number: Z19990
A:Accession: T25169
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-330 <WIL>
A:Cross-references: EMBL:Z81129; PIDN:CAB03405.1; GSPDB:GN00023; CESP:T23F1.6
C:Genetics:
A:Gene: CESP:T23F1.6
A:Map position: 5

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 25, 2002, 16:22:42 ; Search time 27.55 Seconds
(without alignments)
233,301 Million cell updates/sec

Title: US-09-854-864-15

Perfect score: 909

Sequence: 1 MSGLGRSGRRGSRVDQER.....SPALPGLKLSADQVALVYST 166

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	97.5	10.7	1877	1	PCK5_MOUSE
2	83	9.1	1150	1	APWU_PIG
3	82.5	9.1	913	1	PCK5_HUMAN
4	82	9.0	4660	1	LRP2_RAT
5	81	8.9	1124	1	TIE2_HUMAN
6	80.5	8.9	937	1	PAC4_RAT
7	79.5	8.7	1172	1	LMB3_HUMAN
8	79.5	8.7	1984	1	YLDRONE
9	78.5	8.6	630	1	Z147_HUMAN
10	77.5	8.5	2871	1	FBN1_MOUSE
11	76.5	8.4	1680	1	FUR2_DROME
12	76	8.4	603	1	CFAI_MOUSE
13	76	8.4	1877	1	PCK5_RAT
14	75.5	8.3	427	1	YK5_YEAST
15	75	8.3	400	1	LMB_HIRME
16	75	8.3	446	1	FAT_MOUSE
17	75	8.3	969	1	PAC4_HUMAN
18	75	8.3	2871	1	FBN1_MOUSE
19	75	8.3	4544	1	LEP1_HUMAN
20	74.5	8.2	501	1	TRA2_MOUSE
21	74.5	8.2	773	1	ITB1_BOVIN
22	74	8.1	1125	1	TIE2_BOVIN
23	74	8.1	1426	1	EGFR_DROME
24	74	8.1	2569	1	LMA3_MOUSE
25	73.5	8.1	775	1	PTNC_MOUSE
26	73.5	8.1	798	1	ITB1_MOUSE
27	73.5	8.1	1693	1	RIP2_MOUSE
28	73	8.0	61	1	MTID_PIG
29	73	8.0	1339	1	ERB3_RAT
30	73	8.0	4393	1	PGBM_HUMAN
31	72.5	8.0	450	1	NH14_CAEEL
32	72.5	8.0	634	1	Z147_MOUSE
33	72.5	8.0	780	1	PTNC_HUMAN

RESULT 1

ID	PCK5_MOUSE	STANDARD;	PRT;	1877 AA.
AC	Q04592; Q62040;			
DT	01-FEB-1995 (Rel. 31, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Protein convertase subtilisin/kexin type 5 precursor (EC 3.4.21.-)			
DE	(Protein convertase PC5) (Subtilisin/kexin-like protease PC5)			
DE	(Convertase PC5) (PC6) (Subtilisin-like proprotein convertase 6)			
DE	(SPC6).			
GN	PCSK5.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE OF 330-1877 FROM N.A. (ISOFORM PC5B).			
RC	STRAIN=ICR; TISSUE=Intestine;			
RX	MEDLINE=93327934; PubMed=8335106;			
RA	Nakagawa T., Murakami K., Nakayama K.;			
RT	"Identification of an isoform with an extremely large Cys-rich region of PC6, a Kex2-like processing endoprotease.";			
RL	FEBS Lett. 327:165-171(1993).			
RN	[2]			
RP	SEQUENCE FROM N.A. (ISOFORM PC5A).			
RC	TISSUE=Brain, and Intestine;			
RX	MEDLINE=93224489; PubMed=8468318;			
RA	Nakagawa T., Hosaka M., Torii S., Watanabe T., Murakami K.;			
RA	Nakayama K.;			
RT	"Identification and functional expression of a new member of the mammalian Kex2-like processing endoprotease family: its striking structural similarity to PACE4.";			
RL	J. Biochem. 113:132-135(1993).			
RN	[3]			
RP	SEQUENCE FROM N.A. (ISOFORM PC5A).			
RC	TISSUE=Adrenal cortex;			
RX	MEDLINE=93342056; PubMed=8341687;			
RA	Lusson J., Vieus D., Hamelin J., Day R., Chretien M., Seidah N.G.;			
RT	"cDNA structure of the mouse and rat subtilisin/kexin-like PC5: a candidate proprotein convertase expressed in endocrine and nonendocrine cells.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 90:6691-6695(1993).			
RN	[4]			
RP	PARTIAL SEQUENCE, AND SUBCELLULAR LOCATION.			
RX	MEDLINE=97103178; PubMed=8947550;			
RA	De Bie I., Marcinkiewicz M., Malide D., Lazure C., Nakayama K., Bendayan M., Seidah N.G.;			
RT	"The isoforms of proprotein convertase PC5 are sorted to different subcellular compartments.";			
RL	J. Cell Biol. 135:1261-1275(1996).			
RN	[5]			
RP	DEVELOPMENTAL EXPRESSION.			
RX	MEDLINE=96293359; PubMed=8698813;			
RA	Constam D.B., Calton M., Robertson E.J.;			
RT	"SPC4, SPC6, and the novel protease SPC7 are coexpressed with bone			

P18564 homo sapien
Q9nj15 branchiosto
P13378 homo sapien
P36184 entamoeba h
O43915 homo sapien
P50555 sus scrofa
P20736 boophilus m
P20662 mus musculu
O54826 mus musculu
P24348 caenorhabdi
P14585 caenorhabdi
P97927 mus musculu

ALIGNMENTS

34	72.5	8.0	788	1	ITB6_HUMAN
35	72.5	8.0	1696	1	PCK5_BRACL
36	72	7.9	290	1	HXD8_HUMAN
37	72	7.9	308	1	ACPL_ENTHI
38	72	7.9	354	1	VEGD_HUMAN
39	72	7.9	461	1	TRIA_PIG
40	72	7.9	650	1	BM86_BOOMI
41	72	7.9	783	1	ZFY2_MOUSE
42	72	7.9	1068	1	AF1Q_MOUSE
43	72	7.9	1323	1	LT23_CAEEL
44	72	7.9	1429	1	LI12_CAEEL
45	72	7.9	1816	1	LMA4_MOUSE

morphogenetic proteins at distinct sites during embryogenesis.";
J. Cell Biol. 134:181-191(1996).
[6]
RT DEVELOPMENTAL EXPRESSION.
RN MEDLINE=97436919; PubMed=9291583;
RP Rancourt S.L.; Rancourt D.E.;
RA "Murine subtilisin-like proteinase SPC6 is expressed during embryonic
RT implantation, somitogenesis, and skeletal formation";
RL Dev. Genet. 21:75-81(1997).
CC -1- FUNCTION: LIKELY TO REPRESENT A WIDESPREAD ENDOPEPTIDASE ACTIVITY
CC WITHIN THE CONSTITUTIVE AND REGULATED SECRETORY PATHWAY. CAPABLE
CC OF CLEAVAGE AT THE RX(K/R)R CONSENSUS MOTIF. MAY BE RESPONSIBLE
CC FOR THE MATURATION OF GASTROINTESTINAL PEPTIDES. MAY BE INVOLVED
CC IN THE CELLULAR PROLIFERATION OF ADRENAL CORTEX VIA THE ACTIVATION
CC OF GROWTH FACTORS.
CC -1- CATALYTIC ACTIVITY: RELEASE OF MATURE PROTEINS FROM THEIR
CC PROTEINS BY CLEAVAGE OF ARG-XAA-YAA-ARG-|-ZAA BONDS, WHERE XAA
CC CAN BE ANY AMINO ACID AND YAA IS ARG OR LYS.
CC -1- SUBCELLULAR LOCATION: PCSA IS SECRETED THROUGH THE REGULATED
CC SECRETORY PATHWAY. PCSB IS A TYPE I MEMBRANE PROTEIN LOCALIZED TO
CC A PARANUCLEAR POST-GOLGI NETWORK COMPARTMENT IN COMMUNICATION WITH
CC EARLY ENDOSOMES.
CC -1- ALTERNATIVE PRODUCTS: AT LEAST 2 ISOFORMS: PCS5B/LONG (SHOWN HERE)
CC AND PCS5A/SHORT; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: PCSA IS EXPRESSED IN MOST TISSUES BUT IS MOST
CC ABUNDANT IN THE INTESTINE AND ADRENALS. PCSB IS EXPRESSED IN THE
CC INTESTINE, ADRENALS AND LUNG BUT NOT IN THE BRAIN.
CC -1- DEVELOPMENTAL STAGE: WEAKLY EXPRESSED THROUGHOUT THE EMBRYO,
CC BUT MARKEDLY UPREGULATED AT DISCRETE SITES DURING DEVELOPMENT. AT
CC E6.5, PROMINENT EXPRESSION OBSERVED IN DIFFERENTIATED DECIDUA. AT
CC E7.5, INTENSE EXPRESSION IN EXTRAEMBRYONIC ENDODERM, AMNION AND
CC NASCENT MESODERM. AT 8.5, ABUNDANT EXPRESSION IN SOMITES AND YOLK
CC SAC FOLLOWED BY A CONFINATION TO DERMAMYOTOME COMPARTMENT. BETWEEN
CC E9.5 AND E11.5, ABUNDANT EXPRESSION IN AER (THICKENED ECTODERMAL
CC CELLS OF LIMB BUDS). AT E12.5, EXPRESSION IN THE LIMBS IS CONFINED
CC TO THE CONDENSING MESENCHYM SURROUNDING THE CARTILAGE. AT THIS
CC STAGE, STRONG EXPRESSION ALSO DETECTED IN VERTEBRAL AND FACIAL
CC CARTILAGE PRIMORDIA AND IN THE MUSCLE OF THE TONGUE. AT E16.5,
CC ABUNDANT EXPRESSION IN EPITHELIAL CELLS OF THE INTESTINAL VILLI.
CC ISOFORM A IS MOST ABUNDANT AT ALL STAGES BUT SIGNIFICANT LEVELS OF
CC ISOFORM B OCCUR AT E12.5.
CC -1- DOMAIN: THE PROPEPTIDE DOMAIN ACTS AS AN INTRAMOLECULAR CHAPERONE
CC ASSISTING THE FOLDING OF THE ZYMOGEN WITHIN THE ENDOPLASMIC
CC RETICULUM.
CC -1- DOMAIN: AC 1 AND AC 2 (CLUSTERS OF ACIDIC AMINO ACIDS) CONTAIN
CC SORTING INFORMATION. AC 1 DIRECTS TGN LOCALIZATION AND INTERACTS
CC WITH THE TGN SORTING PROTEIN PAGA-1.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8; ALSO KNOWN AS THE
CC SUBTILASE FAMILY.
CC -1- SIMILARITY: CONTAINS 1 HOMO B/P DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
CC or send an email to license@isb-sib.ch.
CC -----
CC EMBL; D17583; BAA04507.1; -;
CC EMBL; D12619; BAA02143.1; -;
CC EMBL; L14932; AAA74636.1; -;
CC PIR; JX0248; JX0248.
CC PIR; A48225; A48225.
CC HSP; Q99405; IMPT.
CC MROPS; S08.076; -;
CC MGD; MGI:97515; Pcsk5.
CC InterPro; IPR000561; EGF-like.
CC InterPro; IPR002174; Furin-like.
CC InterPro; IPR002884; P_domain.
CC InterPro; IPR000209; Peptidase_S8.
CC Pfam; PF01483; P; 1.

RESULT 3
PK55_HUMAN
ID PK55_HUMAN STANDARD; PRT; 913 AA.
AC Q92824; Q13527;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Proprotein convertase subtilisin/kexin type 5 precursor (EC 3.4.21.-)
DE (Proteinase convertase PC5) (Subtilisin/kexin-like protease PC5)
DE (Convertase PC5) (PC6) (hpc6).
GN PCSK5 OR PC5 OR PC6.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=T-cell;
RX MEDLINE=9635380; PubMed=8755538;
RA Miranda L., Wolf J., Pichuanes S., Duke R., Franzusoff A.;
RT "Isolation of the human PC6 gene encoding the putative host protease
for HIV-1 gp160 processing in CD4+ T lymphocytes.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:7695-7700(1996).
RN [2]
RP REVISIONS.
RA Franzusoff A., Miranda L., Wolf J., Pichuanes S., Lu Y., Duke R.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 15-913 FROM N.A.
RA Reudelhuber T.L.;
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: LIKELY TO REPRESENT A WIDESPREAD ENDOPEPTIDASE ACTIVITY
CC WITHIN THE CONSTITUTIVE AND REGULATED SECRETORY PATHWAY. CAPABLE
CC OF CLEAVAGE AT THE RX(K/R)R CONSENSUS MOTIF.
CC -!- CATALYTIC ACTIVITY: RELEASE OF MATURE PROTEINS FROM THEIR
CC PROPEPTIDES BY CLEAVAGE OF ARG-YAA-YAA-ARG-|-ZAA BONDS, WHERE XAA
CC CAN BE ANY AMINO ACID AND YAA IS ARG OR LYS.
CC -!- SUBCELLULAR LOCATION: Secreted (by similarity).
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS SEEMS TO BE PRODUCED BY
CC ALTERNATIVE SPLICING.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN T-LYMPHOCYTES.
CC -!- DOMAIN: THE PROPEPTIDE DOMAIN ACTS AS AN INTRAMOLECULAR CHAPERONE
CC ASSISTING THE FOLDING OF THE ZYMOGEN WITHIN THE ENDOPLASMIC
CC RETICULUM.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8; ALSO KNOWN AS THE
CC SUBTILASE FAMILY.
CC -!- SIMILARITY: CONTAINS 1 HOMO B/P DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U56387; AAC50643.2; -;
DR EMBL; U49114; AAA91807.1; -;
DR HSSP; Q99405; 1MPT.
DR MIM; 600488; -;
DR InterPro; IPR002174; Furin-like.
DR InterPro; IPR002884; P_domain.
DR InterPro; IPR000209; Peptidase_S8.
DR Pfam; PF01483; P; 1.
DR Pfam; PF00082; Peptidase_S8; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR ProDom; PD000717; P_domain; 1.
DR SMART; SM00261; FU; 5.
DR PROSITE; PS00136; SUBTILASE_ASP; 1.
DR PROSITE; PS00137; SUBTILASE_HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
KW Hydrolase; Serine protease; Glycoprotein; Zymogen; Signal;

KW Cleavage on pair of basic residues; Repeat.
FT SIGNAL 1 32
FT PROPEP 33 114 BY SIMILARITY.
FT CHAIN 115 913 PROPROTEIN CONVERTASE SUBTILISIN/KEXIN
FT TYPE 5.
FT CATALYTIC.
FT DOMAIN 115 454
FT DOMAIN 462 600 HOMO B.
FT DOMAIN 636 913
FT SITE 114 115 CYS-RICH MOTIF (CRM) REGION.
FT SITE 519 521 CLEAVAGE (AUTO-) (BY SIMILARITY).
FT ACT_SITE 171 171 CELL ATTACHMENT SITE (POTENTIAL).
FT ACT_SITE 212 212 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 386 386 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT CARBOHYD 225 225 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT CARBOHYD 381 381 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 665 665 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 752 752 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 802 802 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 852 852 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 118 118 S -> F (IN REF. 3).
FT CONFLICT 121 121 V -> A (IN REF. 3).
FT CONFLICT 511 511 R -> A (IN REF. 3).
FT CONFLICT 601 601 R -> O (IN REF. 3).
SQ SEQUENCE 913 AA; 101775 MW; 21389264CAD7546C CRC64;

Query Match 9.1%; Score 82.5; DB 1; Length 913;
Best Local Similarity 26.0%; Pred. No. 2.4;
Matches 27; Conservative 10; Mismatches 34; Indels 33; Gaps 7;

Qy 27 TGVAMRSCPEQYWDPLLGTCMSCKTICNHOSQRTCAAF-----CR-----SLSC 71
Db 717 TNSCVTHCPDGSQDTKNLCRCSENC-----KTCTEFHNCETCRDGLSLQSGRCSVSC 771
Qy 72 RKEGKGYDHLRLRCISCASCISIC-----GQHPKQC-----AYFCEN 106
Db 772 --EDGRFPNG--QDCQPCHRFCATCAGAGAGCGCINCIEGYFWEED 811

RESULT 4
LRP2_RAT
ID LRP2_RAT STANDARD; PRT; 4660 AA.
AC P98158;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Low-density lipoprotein receptor-related protein 2 precursor (Megalin)
DE (Glycoprotein 330).
GN LRP2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Kidney;
RX MEDLINE=95024033; PubMed=7937880;
RA Saito A., Pietromonaco S., Loo A.K.C., Farquhar M.G.;
RT "Complete cloning and sequencing of rat gp330/'megalin,' a
RT distinctive member of the low density lipoprotein receptor gene
RT family.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:9725-9729(1994).
RN [2]
RP FUNCTION.
RX MEDLINE=95386696; PubMed=7544804;
RA Moestrup S.K., Cui S., Vorum H., Bregengaard C., Bjorn S.E.,
RA Norris K., Gliemann J., Christensen E.T.;
RT "Evidence that epithelial glycoprotein 330/megalin mediates uptake of
RT polybasic drugs.";
RL J. Clin. Invest. 96:1404-1413(1995).
RN [3]
RP TISSUE SPECIFICITY.
RX MEDLINE=94172242; PubMed=7510321;

RA Zheng G., Bachinsky D.R., Stamenkovic I., Strickland D.K., Brown D.,
 RA Andres G., McCluskey R.T.;
 RA "Organ distribution in rats of two members of the low-density
 RT lipoprotein receptor gene family, gp330 and Lrp/alpha 2MR, and the
 RT receptor-associated protein (RAP)," *J. Biol. Chem.* 269:11442-11448 (1994).
 RL J. Histochem. Cytochem. 42:531-542(1994).
 CC -1- FUNCTION: BINDS PLASMINOGEN, EXTRACELLULAR MATRIX COMPONENTS,
 CC COMPLEX, APOLOPROTEIN B-ENRICHED ACTIVATOR INHIBITOR TYPE I
 CC LACTOFERRIN, CLUSTERIN AND CALCIUM.
 CC -1- FUNCTION: RECEPTOR-MEDIATED UPTAKE OF POLYBASIC DRUGS SUCH AS
 CC APROFININ, AMINOGLYCOSIDES AND POLYMYXIN B.
 CC -1- SUBUNIT: FORMS A MULTIMERIC COMPLEX TOGETHER WITH A RECEPTOR-
 CC ASSOCIATED PROTEIN (RAP).
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. EXPRESSED IN
 CC CLATHRIN-COATED PITS; A SOLUBLE FORM IS POSSIBLY DERIVED BY
 CC CLEAVAGE AT THE CELL SURFACE.
 CC -1- TISSUE SPECIFICITY: EPITHELIAL CELLS OF KIDNEY GLOMERULUS AND
 CC PROXIMAL TUBULE, LUNG, EPIDIDYMIS, YOLK SAC, AMONG OTHERS.
 CC -1- SIMILARITY: CONTAINS 36 LDL-RECEPTOR CLASS A DOMAINS.
 CC -1- SIMILARITY: CONTAINS 37 LDL-RECEPTOR CLASS B DOMAINS.
 CC -1- SIMILARITY: CONTAINS 17 EGF-LIKE DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; L34049; AAA51369.1; -;
 DR HSP; Q07954; ICR8.
 DR GlycoSuiteDB; P98158; -;
 DR InterPro; IPR000152; Asx_hydroxyl.
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR001881; EGF_CA.
 DR InterPro; IPR002172; LDL_receptor_A.
 DR InterPro; IPR000033; LDL_receptor_rep.
 DR Pfam; PF00057; ldl_recept_a; 36.
 DR Pfam; PF00058; ldl_recept_b; 33.
 DR PRINTS; PR00261; LDLRECEPTOR.
 DR SMART; SM00179; EGF_CA; 3.
 DR SMART; SM00001; EGF_Like; 15.
 DR SMART; SM00192; LDLA; 36.
 DR SMART; SM00135; LY; 35.
 DR PROSITE; PS00101; ASX_HYDROXYL; 4.
 DR PROSITE; PS00022; EGF_1; 1.
 DR PROSITE; PS01185; EGF_2; 8.
 DR PROSITE; PS01187; EGF_CA; 3.
 DR PROSITE; PS01209; LDLRA_1; 31.
 DR PROSITE; PS00068; LDLRA_2; 36.
 KW Glycoprotein; Repeat; Endocytosis; Coated pits; Transmembrane;
 KW Receptor; EGF-like domain; SH3-binding; Signal.
 FT SIGNAL 1 25 POTENTIAL..
 FT CHAIN 26 4660 LOW-DENSITY LIPOPROTEIN RECEPTOR-RELATED
 FT PROTEIN 2.
 FT DOMAIN 26 4425 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 4426 4446 POTENTIAL.
 FT DOMAIN 4447 4660 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 26 64 LDL-RECEPTOR CLASS A 1.
 FT DOMAIN 65 105 LDL-RECEPTOR CLASS A 2.
 FT DOMAIN 106 144 LDL-RECEPTOR CLASS A 3.
 FT DOMAIN 145 181 LDL-RECEPTOR CLASS A 4.
 FT DOMAIN 181 219 LDL-RECEPTOR CLASS A 5.
 FT DOMAIN 220 258 LDL-RECEPTOR CLASS A 6.
 FT DOMAIN 263 307 LDL-RECEPTOR CLASS A 7.
 FT DOMAIN 308 346 EGF-LIKE 1.
 FT DOMAIN 347 385 EGF-LIKE 2.
 FT DOMAIN 435 476 LDL-RECEPTOR CLASS B 1.
 FT DOMAIN 478 519 LDL-RECEPTOR CLASS B 2.
 FT DOMAIN 521 566 LDL-RECEPTOR CLASS B 3.
 FT DOMAIN 568 611 LDL-RECEPTOR CLASS B 4.

DOMAIN 612 652 LDL-RECEPTOR CLASS B 5.
 DOMAIN 658 704 EGF-LIKE 3.
 DOMAIN 752 793 LDL-RECEPTOR CLASS B 6.
 DOMAIN 795 835 LDL-RECEPTOR CLASS B 7.
 DOMAIN 837 879 LDL-RECEPTOR CLASS B 8.
 DOMAIN 881 923 LDL-RECEPTOR CLASS B 9.
 DOMAIN 969 1013 EGF-LIKE 4.
 DOMAIN 1023 1061 LDL-RECEPTOR CLASS A 8.
 DOMAIN 1064 1103 LDL-RECEPTOR CLASS A 9.
 DOMAIN 1108 1146 LDL-RECEPTOR CLASS A 10.
 DOMAIN 1148 1186 LDL-RECEPTOR CLASS A 11.
 DOMAIN 1186 1225 LDL-RECEPTOR CLASS A 12.
 DOMAIN 1229 1269 LDL-RECEPTOR CLASS A 13.
 DOMAIN 1270 1308 LDL-RECEPTOR CLASS A 14.
 DOMAIN 1311 1351 LDL-RECEPTOR CLASS A 15.
 DOMAIN 1350 1390 EGF-LIKE 5.
 DOMAIN 1391 1430 EGF-LIKE 6.
 DOMAIN 1479 1520 LDL-RECEPTOR CLASS B 10.
 DOMAIN 1522 1563 LDL-RECEPTOR CLASS B 11.
 DOMAIN 1567 1609 LDL-RECEPTOR CLASS B 12.
 DOMAIN 1611 1654 LDL-RECEPTOR CLASS B 13.
 DOMAIN 1656 1696 LDL-RECEPTOR CLASS B 14.
 DOMAIN 1701 1742 EGF-LIKE 7.
 DOMAIN 1791 1832 LDL-RECEPTOR CLASS B 15.
 DOMAIN 1834 1882 LDL-RECEPTOR CLASS B 16.
 DOMAIN 1884 1930 LDL-RECEPTOR CLASS B 17.
 DOMAIN 1932 1972 LDL-RECEPTOR CLASS B 18.
 DOMAIN 1973 2013 LDL-RECEPTOR CLASS B 19.
 DOMAIN 2019 2060 EGF-LIKE 8.
 DOMAIN 2108 2156 LDL-RECEPTOR CLASS B 20.
 DOMAIN 2158 2201 LDL-RECEPTOR CLASS B 21.
 DOMAIN 2203 2245 LDL-RECEPTOR CLASS B 22.
 DOMAIN 2247 2289 LDL-RECEPTOR CLASS B 23.
 DOMAIN 2291 2332 LDL-RECEPTOR CLASS B 24.
 DOMAIN 2343 2384 EGF-LIKE 9.
 DOMAIN 2432 2477 LDL-RECEPTOR CLASS B 25.
 DOMAIN 2479 2518 LDL-RECEPTOR CLASS B 26.
 DOMAIN 2520 2562 LDL-RECEPTOR CLASS B 27.
 DOMAIN 2564 2604 LDL-RECEPTOR CLASS B 28.
 DOMAIN 2605 2647 LDL-RECEPTOR CLASS B 29.
 DOMAIN 2652 2694 EGF-LIKE 10.
 DOMAIN 2699 2739 LDL-RECEPTOR CLASS A 16.
 DOMAIN 2740 2778 LDL-RECEPTOR CLASS A 17.
 DOMAIN 2779 2820 LDL-RECEPTOR CLASS A 18.
 DOMAIN 2821 2862 LDL-RECEPTOR CLASS A 19.
 DOMAIN 2863 2903 LDL-RECEPTOR CLASS A 20.
 DOMAIN 2906 2947 LDL-RECEPTOR CLASS A 21.
 DOMAIN 2948 2992 LDL-RECEPTOR CLASS A 22.
 DOMAIN 2993 3031 LDL-RECEPTOR CLASS A 23.
 DOMAIN 3032 3072 LDL-RECEPTOR CLASS A 24.
 DOMAIN 3075 3112 LDL-RECEPTOR CLASS A 25.
 DOMAIN 3113 3153 EGF-LIKE 11.
 DOMAIN 3154 3194 EGF-LIKE 12.
 DOMAIN 3241 3282 LDL-RECEPTOR CLASS B 10.
 DOMAIN 3284 3333 LDL-RECEPTOR CLASS B 11.
 DOMAIN 3335 3377 LDL-RECEPTOR CLASS B 12.
 DOMAIN 3379 3420 LDL-RECEPTOR CLASS B 13.
 DOMAIN 3421 3461 LDL-RECEPTOR CLASS B 14.
 DOMAIN 3467 3511 EGF-LIKE 13.
 DOMAIN 3512 3552 LDL-RECEPTOR CLASS A 26.
 DOMAIN 3553 3593 LDL-RECEPTOR CLASS A 27.
 DOMAIN 3594 3634 LDL-RECEPTOR CLASS A 28.
 DOMAIN 3635 3675 LDL-RECEPTOR CLASS A 29.
 DOMAIN 3678 3718 LDL-RECEPTOR CLASS A 30.
 DOMAIN 3719 3758 LDL-RECEPTOR CLASS A 31.
 DOMAIN 3759 3797 LDL-RECEPTOR CLASS A 32.
 DOMAIN 3798 3836 LDL-RECEPTOR CLASS A 33.
 DOMAIN 3842 3882 LDL-RECEPTOR CLASS A 34.
 DOMAIN 3883 3924 LDL-RECEPTOR CLASS A 35.
 DOMAIN 3928 3966 LDL-RECEPTOR CLASS A 36.
 DOMAIN 3968 4008 EGF-LIKE 14.
 DOMAIN 4009 4050 EGF-LIKE 15.
 DOMAIN 4156 4197 LDL-RECEPTOR CLASS B 35.

```

FT DOMAIN 4199 4241 LDL-RECEPTOR CLASS B 36.
FT DOMAIN 4244 4284 LDL-RECEPTOR CLASS B 37.
FT DOMAIN 4332 4370 EGF-LIKE 16.
FT DOMAIN 4379 4413 EGF-LIKE 17.
FT SITE 4454 4460 SH3-BINDING (POTENTIAL).
FT SITE 4457 4463 SH3-BINDING (POTENTIAL).
FT SITE 4606 4609 SH2-BINDING (POTENTIAL).
FT SITE 4619 4625 SH2-BINDING (POTENTIAL).
FT SITE 4624 4630 SH3-BINDING (POTENTIAL).
FT SITE 1743 1745 CELL ATTACHMENT SITE (POTENTIAL).
FT SITE 4522 4527 ENDOCYTOSIS SIGNAL (POTENTIAL).
FT SITE 4601 4606 ENDOCYTOSIS SIGNAL (POTENTIAL).
FT DISULFID 28 40 BY SIMILARITY.
FT DISULFID 35 53 BY SIMILARITY.
FT DISULFID 47 62 BY SIMILARITY.
FT DISULFID 67 80 BY SIMILARITY.
FT DISULFID 74 93 BY SIMILARITY.
FT DISULFID 87 103 BY SIMILARITY.
FT DISULFID 108 120 BY SIMILARITY.
FT DISULFID 115 133 BY SIMILARITY.
FT DISULFID 127 142 BY SIMILARITY.
FT DISULFID 147 157 BY SIMILARITY.

Query Match 9.0%; Score 82; DB 1; Length 4660;
Best Local Similarity 25.8%; Pred. No. 13;
Matches 40; Conservative 15; Mismatches 64; Indels 36; Gaps 8;

QY 33 SCPEQY-WDPLGTCMSCKTTCNHQ---SORTCAAFRCSLSCREQKGFYDH----- 81
Db 3076 TCPLQHFRCND--GHCIEMGRVNCNVHDCSDNSDEKGGINECLDSSISRCDHNTDIT 3133
QY 82 -----LLRDCISCAST--CGQHPKQCAFCENKLSPV-NLPLELRQRSGEVE 127
Db 3134 SPYCSCLPGYKLMSDKRSCVIDECKESPQLCSQKQCNVSVYICKAPGYIREPDGKSC 3193
QY 128 NNSDN-----SGRY--QGLEHRCSEASPALGL 153
Db 3194 RONSNIERYLFSNRYRNLTTDGSYSLLIQL 3228

RESULT 5
ID TIE2_HUMAN STANDARD; PRT; 1124 AA.
AC O02763.
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Angiopoietin 1 receptor precursor (EC 2.7.1.112) (Tyrosine-protein
DE kinase receptor TIE-2) (Tyrosine-protein kinase receptor TEK) (P140
DE TEK) (Tunica interna endothelial cell kinase).
GN TEK OR TIE2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=93173509; PubMed=8382358;
RA Ziegler S.F., Bird T.A., Schneringer J.A., Schooley K.A., Baum P.R.;
RT "Molecular cloning and characterization of a novel receptor protein
RT tyrosine kinase from human placenta.";
RL Oncogene 8:663-670(1993).
[2]
RN VARIANT VMC1 TRP-849.
RX MEDLINE=97134665; PubMed=8980225;
RA Vekula M., Boon L.M., Carraway K.L. III, Calvert J.T., Diamonti A.J.,
RA Goumnerov B., Pasyk K.A., Marchuk D.A., Warman M.L., Cantley L.C.,
RA Mulliken J.B., Olse B.R.;
RT "Vascular dysmorphogenesis caused by an activating mutation in the
RT receptor tyrosine kinase TIE2.";
RL Cell 87:1181-1190(1996).
[3]

```

```

RP VARIANTS VMC1 TRP-849 AND SER-897.
RX MEDLINE=99299243; PubMed=10369874;
RA Calvert J.T., Riney T.J., Kontos C.D., Cha E.H., Prieto V.G.,
RA Shea C.R., Berg J.N., Nevin N.C., Simpson S.A., Pasyk K.A.,
RA Speer M.C., Peters K.G., Marchuk D.A.;
RT "Allelic and locus heterogeneity in inherited venous malformations.";
RL Hum. Mol. Genet. 8:1279-1289(1999).
CC -!- FUNCTION: THIS PROTEIN IS A PROTEIN TYROSINE-KINASE TRANSMEMBRANE
CC RECEPTOR FOR ANGIOPOIETIN 1. IT MAY CONSTITUTE THE EARLIEST
CC MAMMALIAN ENDOTHELIAL CELL LINEAGE MARKER. PROBABLY REGULATES
CC ENDOTHELIAL CELL PROLIFERATION, DIFFERENTIATION AND GUIDES THE
CC PROPER PATTERNING OF ENDOTHELIAL CELLS DURING BLOOD VESSEL
CC FORMATION.
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine -> ADP + protein
CC tyrosine phosphate.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN ENDOTHELIAL CELLS
CC AND THEIR PROGENITORS, THE ANGIOBLASTS. HAS BEEN DIRECTLY FOUND
CC IN PLACENTA AND LUNG, WITH A LOWER LEVEL IN UMBILICAL VEIN
CC ENDOTHELIAL CELLS, BRAIN AND KIDNEY.
CC -!- DISEASE: DEFECTS IN TEK ARE A CAUSE OF DOMINANTLY INHERITED VENOUS
CC MALFORMATIONS (VMC1). AN ERROR OF VASCULAR MORPHOGENESIS
CC CHARACTERIZED BY DILATED, SERPIGINOUS CHANNELS.
CC -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC DOMAIN.
CC -!- SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC -!- SIMILARITY: CONTAINS 3 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC -!- SIMILARITY: CONTAINS 3 EGF-LIKE DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L06139; AAA61139.1; -.
CC HSP; P11362; IFGK.
CC MIM; 600221; -.
CC MIM; 600195; -.
CC InterPro; IPR000561; EGF-like.
CC InterPro; IPR000719; Euk_kinase.
CC InterPro; IPR003961; FN_III.
CC InterPro; IPR001245; Tyr_kinase.
CC Pfam; PF00008; EGF; 1.
CC Pfam; PF00041; fn3; 3.
CC Pfam; PF00069; pkinase; 1.
CC PRINTS; PR00109; TYRKINASE.
CC SMART; SM00181; EGF; 2.
CC SMART; SM00001; EGF_like; 1.
CC SMART; SM00060; FN3; 3.
CC SMART; SM00219; TyrKc; 1.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
CC PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
CC PROSITE; PS00022; EGF_1; 3.
CC PROSITE; PS01186; EGF_2; 3.
CC Repeat; EGF-like domain; Transmembrane; Immunoglobulin domain;
CC Glycoprotein; Phosphorylation; Multigene family; Disease mutation.
FT SIGNAL 1 18
FT CHAIN 19 1124 ANGIOPOIETIN 1 RECEPTOR.
FT DOMAIN 19 745 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 746 770 POTENTIAL.
FT DOMAIN 771 1124 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 44 102 IG-LIKE C2-TYPE DOMAIN 1.
FT DOMAIN 210 252 EGF-LIKE 1.
FT DOMAIN 254 299 EGF-LIKE 2.
FT DOMAIN 301 341 EGF-LIKE 3.
FT DOMAIN 370 424 IG-LIKE C2-TYPE DOMAIN 2.
FT DOMAIN 444 536 FIBRONECTIN TYPE-III 1.
FT DOMAIN 541 634 FIBRONECTIN TYPE-III 2.

```


OS Homo sapiens (Human).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95293372; PubMed=7774918;
RA Pulkkinen L., Gerecke D.R., Christiano A.M., Wagman D.W.,
RT Burgeson R.E., Uitto J.;
RT "Cloning of the beta 3 chain gene (LAMB3) of human laminin 5, a
RT candidate gene in junctional epidermolysis bullosa.";
RL Genomics 25:192-198(1995).
RN [2]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 172-190.
RX MEDLINE=94209274; PubMed=7512558;
RA Gerecke D.R., Wagman D.W., Champilaud M.F., Burgeson R.E.;
RT "The complete primary structure for a novel laminin chain, the
RT laminin Bk chain.";
RL J. Biol. Chem. 269:11073-11080(1994).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=21221101; PubMed=11296269;
RA Robbins P.B., Lin Q., Goodnough J.B., Tian H., Chen X., Khavari P.A.;
RT "In vivo restoration of laminin 5 beta 3 expression and function in
RT junctional epidermolysis bullosa.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:5193-5198(2001).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=96015057; PubMed=9530036;
RA Morishima Y., Ariyama T., Yamanishi K., Abe T., Ueda E., Yasuno H.,
RA Inazawa J.;
RT "Chromosomal loci of 50 human keratinocyte cDNAs assigned by
RT fluorescence in situ hybridization.";
RL Genomics 28:273-279(1995).
RN [5]
RP SEQUENCE FROM N.A.
RA Grafham D.;
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
RN [6]
RP VARIANT JEB LEU-679.
RX MEDLINE=96055522; PubMed=7550237;
RA Pulkkinen L., McGrath J.A., Christiano A.M., Uitto J.;
RT "Detection of sequence variants in the gene encoding the beta 3 chain
RT of laminin 5 (LAMB3).";
RL Hum. Mutat. 6:77-84(1995).
RN [7]
RP VARIANT GABE LYS-210.
RX MEDLINE=99068967; PubMed=9767254;
RA Mellerio J.E., Eady R.A., Atherton D.J., Lake B.D., McGrath J.A.;
RT "E210K mutation in the gene encoding the beta3 chain of laminin-5
RT (LAMB3) is predictive of a phenotype of generalized atrophic benign
RT epidermolysis bullosa.";
RL Br. J. Dermatol. 139:325-331(1998).
CC [1]- FUNCTION: BINDING TO CELLS VIA A HIGH AFFINITY RECEPTOR, LAMININ
CC IS THOUGHT TO MEDIATE THE ATTACHMENT, MIGRATION, & ORGANIZATION OF
CC CELLS INTO TISSUES DURING EMBRYONIC DEVELOPMENT BY INTERACTING
CC WITH OTHER EXTRACELLULAR MATRIX COMPONENTS.
CC [1]- SUBUNIT: LAMININ IS A COMPLEX GLYCOPROTEIN, CONSISTING OF THREE
CC DIFFERENT POLYPEPTIDE CHAINS (ALPHA, BETA, GAMMA), WHICH ARE BOUND
CC TO EACH OTHER BY DISULFIDE BONDS INTO A CROSS-SHAPED MOLECULE
CC COMPRISING ONE LONG & THREE SHORT ARMS WITH GLOBULES AT EACH END.
CC THE BETA-3 CHAIN IS A SUBUNIT OF LAMININ-5 (EPILGRIN/KALININ/
CC NICEIN).
CC [1]- SUBCELLULAR LOCATION: Extracellular.
CC [1]- TISSUE SPECIFICITY: FOUND IN THE BASEMENT MEMBRANES (MAJOR
CC COMPONENT).
CC [1]- DOMAIN: THE ALPHA-HELICAL DOMAINS I AND II ARE THOUGHT TO INTERACT
CC WITH OTHER LAMININ CHAINS TO FORM A COILED COIL STRUCTURE.
CC [1]- DOMAIN: DOMAIN VI IS GLOBULAR.
CC [1]- DISEASE: DEFECTS IN LAMB3 ARE A CAUSE OF JUNCTIONAL EPIDERMOLYSIS
CC BULLOSA (JEB) GRAVIS, ALSO KNOWN AS HERLITZ TYPE JEB. IT IS A
CC BLISTERING DISORDER IN SKIN THAT IS CHARACTERIZED BY A SEPARATION
CC OF BASAL CELLS FROM THE BASEMENT MEMBRANE DUE TO A DECREASED
CC NUMBER OF HEMIDESMOSOMES. LAMININ-5 IS MISSING FROM THE BASEMENT
CC MEMBRANE OF PATIENTS WITH THE GRAVIS FORM OF EPIDERMOLYSIS
CC BULLOSA.
CC [1]- DISEASE: DEFECTS IN LAMB3 ARE A CAUSE OF GENERALIZED ATROPHIC
CC BENIGN EPIDERMOLYSIS BULLOSA (GABEB). THIS NONLETHAL FORM OF
CC JUNCTIONAL EPIDERMOLYSIS BULLOSA IS CHARACTERIZED BY LIFE-LONG
CC BLISTERING OF THE SKIN, ASSOCIATED WITH HAIR AND TOOTH
CC ABNORMALITIES.
CC [1]- SIMILARITY: CONTAINS 1 LAMININ N-TERMINAL DOMAIN (DOMAIN VI).
CC [1]- SIMILARITY: CONTAINS 6 LAMININ EGF-LIKE DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U17760; AAC51352.1; -
DR EMBL; U17745; AAC51352.1; JOINED.
DR EMBL; U17746; AAC51352.1; JOINED.
DR EMBL; U17747; AAC51352.1; JOINED.
DR EMBL; U17748; AAC51352.1; JOINED.
DR EMBL; U17749; AAC51352.1; JOINED.
DR EMBL; U17750; AAC51352.1; JOINED.
DR EMBL; U17751; AAC51352.1; JOINED.
DR EMBL; U17752; AAC51352.1; JOINED.
DR EMBL; U17753; AAC51352.1; JOINED.
DR EMBL; U17754; AAC51352.1; JOINED.
DR EMBL; U17755; AAC51352.1; JOINED.
DR EMBL; U17756; AAC51352.1; JOINED.
DR EMBL; U17757; AAC51352.1; JOINED.
DR EMBL; U17758; AAC51352.1; JOINED.
DR EMBL; U17759; AAC51352.1; JOINED.
DR EMBL; L25541; AAA61834.1; -
DR EMBL; AY035783; AAK61364.1; -
DR EMBL; D37766; BAA22263.1; -
DR EMBL; AL031316; CAB40149.1; -
DR EMBL; AL023754; CAA19297.1; -
DR HSSP; P02468; IKLO.
DR MM; 150310; -
DR MM; 226650; -
DR MM; 226700; -
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001886; LamNT.
DR InterPro; IPR002049; Laminin_EGF.
DR Pfam; PF00053; laminin_EGF; 6.
DR Pfam; PF00055; laminin_Nterm; 1.
DR PRINTS; PR00011; EGF_LAMININ.
DR ProDom; PD002082; LamNT; 1.
DR SMART; SM00180; EGF_Lam; 6.
DR SMART; SM00136; LamNT; 1.
DR PROSITE; PS00022; EGF_1; 5.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01248; LAMININ_TYPE_EGF; 5.
DR GlycoProtein; Basement membrane; Extracellular matrix; Coiled coil;
KW Laminin EGF-like domain; Cell adhesion; Repeat; Signal;
KW Disease mutation; Epidermolysis bullosa.
FT SIGNAL 1 17 POTENTIAL.
FT CHAIN 18 1172 LAMININ BETA-3 CHAIN.
FT DOMAIN 18 249 LAMININ N-TERMINAL (DOMAIN VI).
FT DOMAIN 250 315 LAMININ EGF-LIKE 1.
FT DOMAIN 316 378 LAMININ EGF-LIKE 2.
FT DOMAIN 379 430 LAMININ EGF-LIKE 3.
FT DOMAIN 431 480 LAMININ EGF-LIKE 4.
FT DOMAIN 481 533 LAMININ EGF-LIKE 5.
FT DOMAIN 534 578 LAMININ EGF-LIKE 6.
FT DOMAIN 579 785 DOMAIN II.
FT DOMAIN 786 816 DOMAIN ALPHA.
FT DOMAIN 817 1170 DOMAIN I.
FT DOMAIN 723 757 COILED COIL (POTENTIAL).

FT	DOMAIN	1734	1770	EGF-LIKE 7.
FT	SITE	1837	1837	CRITICAL FOR ENDOCYTOSIS (POTENTIAL).
FT	SITE	1878	1878	CRITICAL FOR ENDOCYTOSIS (POTENTIAL).
FT	SITE	1892	1892	CRITICAL FOR ENDOCYTOSIS (POTENTIAL).
FT	DISULFID	90	102	BY SIMILARITY.
FT	DISULFID	97	115	BY SIMILARITY.
FT	DISULFID	109	124	BY SIMILARITY.
FT	DISULFID	129	144	BY SIMILARITY.
FT	DISULFID	137	157	BY SIMILARITY.
FT	DISULFID	151	166	BY SIMILARITY.
FT	DISULFID	184	197	BY SIMILARITY.
FT	DISULFID	191	210	BY SIMILARITY.
FT	DISULFID	204	220	BY SIMILARITY.
FT	DISULFID	227	239	BY SIMILARITY.
FT	DISULFID	234	253	BY SIMILARITY.
FT	DISULFID	247	262	BY SIMILARITY.
FT	DISULFID	266	281	BY SIMILARITY.
FT	DISULFID	275	294	BY SIMILARITY.
FT	DISULFID	288	304	BY SIMILARITY.
FT	DISULFID	310	321	BY SIMILARITY.
FT	DISULFID	315	331	BY SIMILARITY.
FT	DISULFID	333	346	BY SIMILARITY.
FT	DISULFID	352	363	BY SIMILARITY.
FT	DISULFID	359	372	BY SIMILARITY.
FT	DISULFID	374	387	BY SIMILARITY.
FT	DISULFID	664	673	BY SIMILARITY.
FT	DISULFID	669	685	BY SIMILARITY.
FT	DISULFID	687	700	BY SIMILARITY.
FT	DISULFID	988	999	BY SIMILARITY.
FT	DISULFID	995	1009	BY SIMILARITY.
FT	DISULFID	1011	1025	BY SIMILARITY.
FT	DISULFID	1031	1040	BY SIMILARITY.
FT	DISULFID	1035	1053	BY SIMILARITY.
FT	DISULFID	1047	1062	BY SIMILARITY.
FT	DISULFID	1074	1087	BY SIMILARITY.
FT	DISULFID	1081	1100	BY SIMILARITY.
FT	DISULFID	1094	1109	BY SIMILARITY.
FT	DISULFID	1118	1130	BY SIMILARITY.
FT	DISULFID	1125	1143	BY SIMILARITY.
FT	DISULFID	1137	1152	BY SIMILARITY.
FT	DISULFID	1158	1170	BY SIMILARITY.
FT	DISULFID	1165	1183	BY SIMILARITY.
FT	DISULFID	1177	1193	BY SIMILARITY.
FT	DISULFID	1198	1210	BY SIMILARITY.
FT	DISULFID	1205	1223	BY SIMILARITY.
FT	DISULFID	1217	1232	BY SIMILARITY.
FT	DISULFID	1243	1257	BY SIMILARITY.
FT	DISULFID	1250	1270	BY SIMILARITY.
FT	DISULFID	1264	1279	BY SIMILARITY.
FT	DISULFID	1283	1296	BY SIMILARITY.
FT	DISULFID	1290	1309	BY SIMILARITY.
FT	DISULFID	1303	1318	BY SIMILARITY.
FT	DISULFID	1340	1352	BY SIMILARITY.
FT	DISULFID	1347	1365	BY SIMILARITY.
FT	DISULFID	1359	1375	BY SIMILARITY.
FT	DISULFID	1379	1392	BY SIMILARITY.
FT	DISULFID	1388	1401	BY SIMILARITY.
FT	DISULFID	1403	1416	BY SIMILARITY.
FT	DISULFID	1422	1432	BY SIMILARITY.
FT	DISULFID	1428	1441	BY SIMILARITY.
FT	DISULFID	1443	1456	BY SIMILARITY.
FT	DISULFID	1738	1747	BY SIMILARITY.
FT	DISULFID	1743	1756	BY SIMILARITY.
FT	DISULFID	1758	1769	BY SIMILARITY.
FT	CARBOHYD	30	30	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	365	365	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	384	384	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	429	429	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	666	666	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	749	749	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	782	782	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1022	1022	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1240	1240	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1265	1265	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1326	1326	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1475	1475	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1490	1490	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	VARIANT	1519	1519	V -> I
SO	SEQUENCE	1984 AA; 219318 MW; F92A256E0ACD5E48	CRC64;	
Query Match 8.7%; Score 79.5; DB 1; Length 1984;				
Best Local Similarity 29.9%; Pred. No. 9.9;				
Matches 32; Conservative 7; Mismatches 41; Indels 27; Gaps 8;				
Qy	26	WTGVAMRSCPEEQYWDPLLTGTCMSCKTICN-----HQSQRTCAAFPCRSLS-----	70	
Db	1333	WS-TSSRSC-RPHLFDCQDGECDLSRVCCNFPDCTNGHDEGPKCATACRSASGRQCQH	1390	
Qy	71	-CRKEQG----KFYD--HLLRDCJSCASI--CGQHPKOCAYFCENKL	108	
Db	1391	KCRATPAGVCSGDFGYRLDADQKSLDIDEC-QEQQPCALCENTL	1436	
RESULT 9				
Z147_HUMAN				
ID	Z147_HUMAN	STANDARD;	PRT;	630 AA.
AC	Q14258;			
DT	16-OCT-2001	(Rel. 40, Created)		
DT	16-OCT-2001	(Rel. 40, Last sequence update)		
DT	01-MAR-2002	(Rel. 41, Last annotation update)		
DE	Zinc finger protein 147 (Estrogen responsive finger protein) (Efp).			
GN	ZNF147 OR EFP.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N. A.			
RC	TISSUE=Placenta;			
RX	MEDLINE=94068555; PubMed=8248217;			
RA	Inoue S., Orimo A., Hosoi T., Kondo S., Toyoshima H., Kondo T.,			
RA	Ikegami A., Ouchi Y., Muramatsu M.;			
RT	"Genomic binding-site cloning reveals an estrogen-responsive gene that			
RT	encodes a RING finger protein."			
RL	Proc. Natl. Acad. Sci. U.S.A. 90:11117-11121(1993).			
CC	-!- FUNCTION: MEDIATES ESTROGEN ACTION IN VARIOUS TARGET ORGANS.			
CC	-!- TISSUE SPECIFICITY: UBIQUITOUS.			
CC	-!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.			
CC	-!- SIMILARITY: CONTAINS 1 SPRY DOMAIN.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; D21205; BAA04747.1;			
DR	MM; 600453;			
DR	InterPro; IPR001870; Gamma_carboxylase.			
DR	InterPro; IPR003877; SPRY.			
DR	InterPro; IPR003878; SPRY_domain.			
DR	InterPro; IPR001841; Znf_ring.			
DR	Pfam; PF00622; SPRY; 1.			
DR	Pfam; PF00097; zf-C3HC4; 1.			
DR	SMART; SM00184; RING; 1.			
DR	SMART; SM00449; SPRY; 1.			
DR	PROSITE; PS00518; ZF_RING_1; 1.			
DR	PROSITE; PS00089; ZF_RING_2; 1.			
KW	Zinc-finger; Coiled coil.			
FT	ZN_FING	13 54		RING-TYPE.
FT	DOMAIN	217 307		COILED COIL (POTENTIAL).
FT	DOMAIN	509 630		SPRY.
FT	SEQUENCE	630 AA; 70989 MW; EE4EE713F96DD080	CRC64;	


```
Query Match      8.6% Score 78.5; DB 1; Length 630;
Best Local Similarity 23.0%; Pred. No. 4;
Matches 38; Conservative 20; Mismatches 54; Indels 53; Gaps 9;

QY 22 PQGLWTGVAMRSPCEQYWDPLLTGTCMSCKTIONH-----SQRTC-----AAFCRS-LSCR 72
   | : || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 86 PADVWTPPARASASP-----NAQVADHCLKEAAVKTCLVCMASFCQEHLOPH 134
   | : || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 73 KEQKGYDHL-----RDCISCASGHPKQCAVFC-----ENKLRSPVNL 115
   | : || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 135 FDSFADHPLQPPVDRLL---RRKCSOHNRLREFFCEHSECHICLVHEKTCSPASL- 191
   | : || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 116 PELRRQSGEVENNSDNGRYQGLEHGRGSEASPALGKLKSADQV 160
   | : || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 192 ----SQASADLEAT-----LHKLTVMYSQINGASRALDDV 223
   | : || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 10
FBN1_MOUSE STANDARD; PRT; 2871 AA.
ID FBN1_MOUSE AC Q61554; Q60826;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Fibrillin 1 precursor.
GN FBN1 OR FBN-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP MEDLINE=95130561; PubMed=7829516;
RA Yin W, Germiller J, Sanguinetti C, Smiley E, Pangillan T,
RA Pereira L, Ramirez F, Bonadio J.;
RT "Primary structure and developmental expression of Fbn-1, the mouse
RL fibrillin gene.";
RL J. Biol. Chem. 270:1798-1806(1995).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN-CD-1; TISSUE-Kidney;
RA Ota K., Kumar A., Wada J., Liu Z., Kanwar Y.S.;
RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: STRUCTURAL COMPONENT OF CONNECTIVE TISSUE MICROFIBRILS
CC THAT BINDS CALCIUM. FIBRILLIN-1-CONTAINING MICROFIBRILS PROVIDE
CC LONG-TERM FORCE BEARING STRUCTURAL SUPPORT.
CC -1- PTM: FORMS INTERMOLECULAR DISULFIDE BONDS EITHER WITH OTHER
CC FIBRILLIN-1 MOLECULES OR WITH OTHER COMPONENTS OF THE
CC MICROFIBRILS (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 43 CALCIUM-BINDING AND 4 NON-CALCIUM BINDING
CC EGF-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 7 TGF-BETA BINDING PROTEIN DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L29454; AAA56840.1; -
CC EMBL; U22493; AAA64217.1; -
CC HSSP; P35555; IAPJ.
CC MGD; MGI:95489; Fbn1.
CC InterPro; IPR000152; Asx_hydroxyl.
CC InterPro; IPR000561; EGF-like.
CC InterPro; IPR001881; EGF_Ca.
CC InterPro; IPR002212; TB.
CC Pfam; PF00008; EGF_46.
CC Pfam; PF00683; TB; 9.

DR SMART; SM00179; EGF_CA; 42.
DR SMART; SM00001; EGF_like; 4.
DR PROSITE; PS0010; ASX_HYDROXYL; 43.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 38.
DR PROSITE; PS01187; EGF_CA; 45.
KW Extracellular matrix; Calcium-binding; Glycoprotein; EGF-like domain;
KW Repeat; Signal; Multigene family.
FT SIGNAL 1 27 POTENTIAL.
FT CHAIN 28 2871 FIBRILLIN 1.
FT DOMAIN 81 112 EGF-LIKE 1, NON-CALCIUM BINDING.
FT DOMAIN 115 146 EGF-LIKE 2, NON-CALCIUM BINDING.
FT DOMAIN 147 178 EGF-LIKE 3, NON-CALCIUM BINDING.
FT DOMAIN 246 287 EGF-LIKE 4, CALCIUM-BINDING.
FT DOMAIN 288 329 EGF-LIKE 5, CALCIUM-BINDING.
FT DOMAIN 330 401 TGFBP 1.
FT DOMAIN 402 446 PRO-RICH.
FT DOMAIN 449 489 EGF-LIKE 6, NON-CALCIUM BINDING.
FT DOMAIN 490 529 EGF-LIKE 7, CALCIUM-BINDING.
FT DOMAIN 530 571 EGF-LIKE 8, CALCIUM-BINDING.
FT DOMAIN 572 612 EGF-LIKE 9, CALCIUM-BINDING.
FT DOMAIN 613 653 EGF-LIKE 10, CALCIUM-BINDING.
FT DOMAIN 656 721 TGFBP 2.
FT DOMAIN 723 764 EGF-LIKE 11, CALCIUM-BINDING.
FT DOMAIN 765 806 EGF-LIKE 12, CALCIUM-BINDING.
FT DOMAIN 807 846 EGF-LIKE 13, CALCIUM-BINDING.
FT DOMAIN 910 951 EGF-LIKE 14, CALCIUM-BINDING.
FT DOMAIN 952 1018 TGFBP 3.
FT DOMAIN 1028 1069 EGF-LIKE 15, CALCIUM-BINDING.
FT DOMAIN 1070 1112 EGF-LIKE 16, CALCIUM-BINDING.
FT DOMAIN 1113 1154 EGF-LIKE 17, CALCIUM-BINDING.
FT DOMAIN 1155 1196 EGF-LIKE 18, CALCIUM-BINDING.
FT DOMAIN 1197 1237 EGF-LIKE 19, CALCIUM-BINDING.
FT DOMAIN 1238 1279 EGF-LIKE 20, CALCIUM-BINDING.
FT DOMAIN 1280 1321 EGF-LIKE 21, CALCIUM-BINDING.
FT DOMAIN 1322 1362 EGF-LIKE 22, CALCIUM-BINDING.
FT DOMAIN 1363 1403 EGF-LIKE 23, CALCIUM-BINDING.
FT DOMAIN 1404 1445 EGF-LIKE 24, CALCIUM-BINDING.
FT DOMAIN 1446 1486 EGF-LIKE 25, CALCIUM-BINDING.
FT DOMAIN 1487 1527 EGF-LIKE 26, CALCIUM-BINDING.
FT DOMAIN 1528 1599 TGFBP 4.
FT DOMAIN 1606 1647 EGF-LIKE 27, CALCIUM-BINDING.
FT DOMAIN 1648 1688 EGF-LIKE 28, CALCIUM-BINDING.
FT DOMAIN 1689 1758 TGFBP 5.
FT DOMAIN 1766 1807 EGF-LIKE 29, CALCIUM-BINDING.
FT DOMAIN 1808 1848 EGF-LIKE 30, CALCIUM-BINDING.
FT DOMAIN 1849 1890 EGF-LIKE 31, CALCIUM-BINDING.
FT DOMAIN 1891 1929 EGF-LIKE 32, CALCIUM-BINDING.
FT DOMAIN 1930 1972 EGF-LIKE 33, CALCIUM-BINDING.
FT DOMAIN 1973 2012 EGF-LIKE 34, CALCIUM-BINDING.
FT DOMAIN 2013 2054 EGF-LIKE 35, CALCIUM-BINDING.
FT DOMAIN 2055 2121 TGFBP 6.
FT DOMAIN 2127 2165 EGF-LIKE 36, CALCIUM-BINDING.
FT DOMAIN 2166 2205 EGF-LIKE 37, CALCIUM-BINDING.
FT DOMAIN 2206 2246 EGF-LIKE 38, CALCIUM-BINDING.
FT DOMAIN 2247 2290 EGF-LIKE 39, CALCIUM-BINDING.
FT DOMAIN 2291 2332 EGF-LIKE 40, CALCIUM-BINDING.
FT DOMAIN 2333 2400 TGFBP 7.
FT DOMAIN 2402 2443 EGF-LIKE 41, CALCIUM-BINDING.
FT DOMAIN 2444 2484 EGF-LIKE 42, CALCIUM-BINDING.
FT DOMAIN 2485 2523 EGF-LIKE 43, CALCIUM-BINDING.
FT DOMAIN 2524 2566 EGF-LIKE 44, CALCIUM-BINDING.
FT DOMAIN 2567 2607 EGF-LIKE 45, CALCIUM-BINDING.
FT DOMAIN 2607 2647 EGF-LIKE 46, CALCIUM-BINDING.
FT DOMAIN 2648 2687 EGF-LIKE 47, CALCIUM-BINDING.
FT DISULFID 85 94 BY SIMILARITY.
FT DISULFID 89 100 BY SIMILARITY.
FT DISULFID 102 111 BY SIMILARITY.
FT DISULFID 119 129 BY SIMILARITY.
FT DISULFID 123 134 BY SIMILARITY.
FT DISULFID 136 145 BY SIMILARITY.
FT DISULFID 150 160 BY SIMILARITY.
FT DISULFID 154 166 BY SIMILARITY.
```

```
FT DISULFID 168 177 BY SIMILARITY.
FT DISULFID 250 262 BY SIMILARITY.
FT DISULFID 257 271 BY SIMILARITY.
FT DISULFID 273 286 BY SIMILARITY.
FT DISULFID 292 304 BY SIMILARITY.
FT DISULFID 299 313 BY SIMILARITY.
FT DISULFID 315 328 BY SIMILARITY.
FT DISULFID 453 465 BY SIMILARITY.
FT DISULFID 460 474 BY SIMILARITY.
FT DISULFID 476 488 BY SIMILARITY.
FT DISULFID 494 504 BY SIMILARITY.
FT DISULFID 499 513 BY SIMILARITY.
FT DISULFID 515 528 BY SIMILARITY.
FT DISULFID 534 546 BY SIMILARITY.
FT DISULFID 541 555 BY SIMILARITY.
FT DISULFID 557 570 BY SIMILARITY.
FT DISULFID 576 587 BY SIMILARITY.
FT DISULFID 582 596 BY SIMILARITY.
FT DISULFID 598 611 BY SIMILARITY.
FT DISULFID 617 628 BY SIMILARITY.
FT DISULFID 623 637 BY SIMILARITY.
FT DISULFID 639 652 BY SIMILARITY.
FT DISULFID 727 739 BY SIMILARITY.
FT DISULFID 734 748 BY SIMILARITY.
FT DISULFID 750 763 BY SIMILARITY.
FT DISULFID 769 781 BY SIMILARITY.
FT DISULFID 776 790 BY SIMILARITY.
FT DISULFID 792 805 BY SIMILARITY.
FT DISULFID 811 821 BY SIMILARITY.
FT DISULFID 816 830 BY SIMILARITY.
FT DISULFID 832 845 BY SIMILARITY.
FT DISULFID 914 926 BY SIMILARITY.
FT DISULFID 921 935 BY SIMILARITY.
FT DISULFID 937 950 BY SIMILARITY.
FT DISULFID 1032 1044 BY SIMILARITY.
FT DISULFID 1039 1053 BY SIMILARITY.
FT DISULFID 1055 1068 BY SIMILARITY.
FT DISULFID 1074 1086 BY SIMILARITY.
FT DISULFID 1081 1095 BY SIMILARITY.
FT DISULFID 1097 1111 BY SIMILARITY.
FT DISULFID 1117 1129 BY SIMILARITY.
FT DISULFID 1124 1138 BY SIMILARITY.
FT DISULFID 1140 1153 BY SIMILARITY.
FT DISULFID 1159 1171 BY SIMILARITY.
FT DISULFID 1166 1180 BY SIMILARITY.
FT DISULFID 1182 1195 BY SIMILARITY.
FT DISULFID 1201 1212 BY SIMILARITY.
FT DISULFID 1208 1221 BY SIMILARITY.
FT DISULFID 1223 1236 BY SIMILARITY.
FT DISULFID 1242 1254 BY SIMILARITY.
FT DISULFID 1249 1263 BY SIMILARITY.
FT DISULFID 1265 1278 BY SIMILARITY.
FT DISULFID 1284 1296 BY SIMILARITY.
FT DISULFID 1291 1305 BY SIMILARITY.
FT DISULFID 1307 1320 BY SIMILARITY.
FT DISULFID 1326 1339 BY SIMILARITY.
FT DISULFID 1333 1348 BY SIMILARITY.
FT DISULFID 1350 1361 BY SIMILARITY.
FT DISULFID 1367 1380 BY SIMILARITY.
FT DISULFID 1374 1389 BY SIMILARITY.
FT DISULFID 1391 1402 BY SIMILARITY.
FT DISULFID 1408 1420 BY SIMILARITY.
FT DISULFID 1415 1429 BY SIMILARITY.
FT DISULFID 1431 1444 BY SIMILARITY.
FT DISULFID 1450 1461 BY SIMILARITY.
FT DISULFID 1456 1470 BY SIMILARITY.
FT DISULFID 1472 1485 BY SIMILARITY.
FT DISULFID 1491 1502 BY SIMILARITY.
FT DISULFID 1497 1511 BY SIMILARITY.
FT DISULFID 1513 1526 BY SIMILARITY.
FT DISULFID 1610 1622 BY SIMILARITY.
FT DISULFID 1617 1631 BY SIMILARITY.
FT DISULFID 1633 1646 BY SIMILARITY.

FT DISULFID 1652 1663 BY SIMILARITY.
FT DISULFID 1658 1672 BY SIMILARITY.

Query Match 8.5% Score 77.5; DB 1; Length 2871;
Best Local Similarity 22.5%; Pred. No. 22; Indels 63; Gaps 9;
Matches 32; Conservative 14; Mismatches 33;

Qy 1 MSGLGRSR-----GGRSRVDQERFPQGLWTGVAMRSCPERQYWDPLLTGTCMS 49
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2347 MCQIGSSNRNPVTKSECCVSGRG-----W-GLHCEICPFE-----GT-VA 2385
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 50 CKTCNH-----QSQTCAAFCSLSCKEQKQFYDHLRDCI----- 87
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2386 YKLCPLHGRGFMNGADVDECKVIHDCVNGECVNDGRSYH-----CICKGTPTDITG 2439
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 88 -SCASI--CGOHPKOCAYFCEN 106
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2440 TSCVDLNECNOAPRCNPFICKN 2461

RESULT 11
FUR2_DROME STANDARD; PRT; 1680 AA.
AC P30432;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Furin-like protease 2 precursor (EC 3.4.21.75) (Furin 2).
GN FUR2.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92381036; PubMed=1512259;
RA Roebroek A.J.M., Greemers J.W.M., Pauli I.G.L., Kurzik-Dumke U.,
RA Rentrop M., Gateff E.A.F., Leunissen J.A.M., van de Ven W.J.;
RA "Cloning and functional expression of Dfurin2, a subtilisin-like
RT proprotein processing enzyme of Drosophila melanogaster with multiple
RT repeats of a cysteine motif";
RT J. Biol. Chem. 267:17208-17215(1992).
RL -1- FUNCTION: FURIN IS LIKELY TO REPRESENT THE UBIQUITOUS ENDOPEPTIDASE
CC ACTIVITY WITHIN CONSTITUTIVE SECRETORY PATHWAYS AND CAPABLE OF
CC CLEAVAGE AT THE RX(K/R)R CONSENSUS MOTIF (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: Release of mature proteins from their
CC propeptides by cleavage of Arg-Xaa-Yaa-Arg-l-Zaa bonds, where Xaa
CC can be any amino acid and Yaa is Arg or Lys. Releases albumin,
CC complement component C3 and von Willebrand factor from their
CC respective precursors.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8; ALSO KNOWN AS THE
CC SUBTILASE FAMILY. FURIN SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M94375; AAA28551.1;
CC PIR; A43434; A43434.
CC HSP; Q99405; 1MPT.
CC FlyBase; FBgn0004598; Fur2.
CC InterPro; IPR000561; EGF-like.
CC InterPro; IPR002174; Furin-like.
CC InterPro; IPR002884; P_Domain.
CC InterPro; IPR000209; Peptidase_S8.
CC Pfam; PF00757; Furin-like; 1.
CC Pfam; PF01483; P; 1.
CC Pfam; PF00082; Peptidase_S8; 1.
```

```
DR PRINTS; PR00723; SUBTILISIN.
DR ProDom; PD000717; P_domain; 1.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00261; FU; 10.
DR PROSITE; PS00136; SUBTILASE_ASP; 1.
DR PROSITE; PS00137; SUBTILASE_HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
DR Hydrolyase; Serine protease; Glycoprotein; Signal; Transmembrane;
KW Multigene family; Zymogen; Repeat.
FT SIGNAL 1 ?
FT PROPEP 319
FT CHAIN 320 1680
FT ACT_SITE 418 457
FT ACT_SITE 457 457
FT ACT_SITE 638 638
FT DOMAIN 962 1444
FT REPEAT 962 1007
FT REPEAT 1008 1057
FT REPEAT 1058 1104
FT REPEAT 1105 1153
FT REPEAT 1154 1205
FT REPEAT 1206 1254
FT REPEAT 1255 1299
FT REPEAT 1300 1346
FT REPEAT 1347 1393
FT REPEAT 1394 1444
FT TRANSMEM 1508 1532
FT DOMAIN 1533 1680
FT CARBOHYD 3 3
FT CARBOHYD 109 109
FT CARBOHYD 130 130
FT CARBOHYD 203 203
FT CARBOHYD 443 443
FT CARBOHYD 481 481
FT CARBOHYD 928 928
FT CARBOHYD 1061 1061
FT CARBOHYD 1182 1182
FT CARBOHYD 1275 1275
FT CARBOHYD 1278 1278
FT CARBOHYD 1440 1440
FT SEQUENCE 1680 AA; 183599 MW; 0A99CE8770A8E293 CRC64;

Query Match 8.4%; Score 76.5; DB 1; Length 1680;
Best Local Similarity 23.6%; Pred. No. 16;
Matches 33; Conservative 15; Mismatches 63; Indels 29; Gaps 6;

Qy 26 WTGVAMSCPEEQYWDPLLTCTMSCKTICNH-QSORTCAAFCSLSCKR----- 73
Db 1139 WQNKCLISCPDGFYADKKRLCEMPCQEGCKTCTSGVCSCLQNTLTKRDKCIVSGSEG 1198

Qy 74 -EQGKFYDHLRLDCISCASICGQHPKQCAVFCENKL-RSPVNLPPELRRRORSGEVENNSD 131
Db 1199 CSEEFYSQVEGQCRPCASCGS-----CNGPADTCTSCPPNRLLEQSRCV----- 1245

Qy 132 NSGRYQG-LEHRGSEASPAL 150
Db 1246 -SGCRGFFVEAGSLCSPL 1264

RESULT 12
ID CFAL_MOUSE STANDARD; PRT; 603 AA.
AC Q61129; Q9WU07;
DT 16-OCT-2001 (Rel. 40, Created);
DT 16-OCT-2001 (Rel. 40, Last sequence update);
DT 16-OCT-2001 (Rel. 40, Last annotation update);
DE Complement factor 1 precursor (EC 3.4.21.45) (C3B/C4b inactivator).
GN IF OR CFI.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
```

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C;
RX MEDLINE=96175003; PubMed=8604219;
RA Minta J.O., Wong M.J., Kozak C.A., Kunnath-Muglia L.M., Goldberger G.;
RT "CDNA cloning, sequencing and chromosomal assignment of the gene for
RT mouse complement factor I (C3b/C4b inactivator): identification of a
RT species specific divergent segment in factor I.";
RL Mol. Immunol. 33:101-112(1996).
RN [2]
RP SEQUENCE OF 114-334 FROM N.A.
RC STRAIN=129/SV; TISSUE=Kidney;
RX MEDLINE=99220287; PubMed=10204086;
RA Yun Y.-S., Goldberger G., Minta J.O.;
RT "Cloning and characterization of the non-catalytic heavy chain of
RT mouse complement factor I gene: structure comparison with the human
RT homologue.";
RL Biochem. Mol. Biol. Int. 47:493-500(1999).
CC -!- FUNCTION: RESPONSIBLE FOR CLEAVING THE ALPHA-CHAINS OF C4B AND C3B
CC IN THE PRESENCE OF THE COFACTORS C4-BINDING PROTEIN AND FACTOR H
CC RESPECTIVELY.
CC -!- SUBUNIT: HETERODIMER OF A LIGHT AND HEAVY CHAINS LINKED BY
CC DISULFIDE BONDS.
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- TISSUE SPECIFICITY: PLASMA.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
CC -!- SIMILARITY: CONTAINS 1 KAZAL-LIKE DOMAIN.
CC -!- SIMILARITY: CONTAINS 2 LDL-RECEPTOR CLASS A DOMAINS.
CC -!- SIMILARITY: CONTAINS 1 SRCR DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF47810; AAB00438.1; -.
DR EMBL; AF100565; AAD32965.1; -.
DR EMBL; AF100559; AAD32965.1; JOINED.
DR EMBL; AF100560; AAD32965.1; JOINED.
DR EMBL; AF100561; AAD32965.1; JOINED.
DR EMBL; AF100562; AAD32965.1; JOINED.
DR EMBL; AF100563; AAD32965.1; JOINED.
DR EMBL; AF100564; AAD32965.1; JOINED.
DR HSSP; P00750; 1RTF.
DR MEROPS; S01.199; -.
DR MGD; MGI:105937; Cfi.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR003884; FcI_MAC.
DR InterPro; IPR002172; LDL_recept_A.
DR InterPro; IPR001190; SRCR.
DR InterPro; IPR001254; Trypsin.
DR InterPro; IPR002350; kazal.
DR Pfam; PF00050; kazal; 1.
DR Pfam; PF00057; ldl_recept_a; 1.
DR Pfam; PF00530; SRCR; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00261; LDLRECEPTOR.
DR SMART; SM00057; FIMAC; 1.
DR SMART; SM00280; KAZAL; 1.
DR SMART; SM00192; LDLA; 2.
DR SMART; SM00202; SR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS01209; LDLRA_1; 1.
DR PROSITE; PS00068; LDLRA_2; 2.
DR PROSITE; PS00420; SRCR_1; FALSE_NEG.
DR PROSITE; PS50287; SRCR_2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
```

```
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Complement pathway; Plasma; Glycoprotein; Hydrolase; Serine protease;
Repeat; Signal.
FT SIGNAL 1 18 POTENTIAL.
FT CHAIN 19 603 COMPLEMENT FACTOR I.
FT CHAIN 19 356 COMPLEMENT FACTOR I HEAVY CHAIN.
FT CHAIN 361 603 COMPLEMENT FACTOR I LIGHT CHAIN.
FT CHAIN 63 109 KAZAL-LIKE.
FT DOMAIN 117 217 SRCR.
FT DOMAIN 218 262 LDL-RECEPTOR CLASS A 1.
FT DOMAIN 263 299 LDL-RECEPTOR CLASS A 2.
FT DOMAIN 361 603 SERINE PROTEASE.
FT DOMAIN 234 252 BY SIMILARITY.
FT DISULFID 246 ?
FT DISULFID 246 276 BY SIMILARITY.
FT DISULFID 271 289 BY SIMILARITY.
FT DISULFID 283 298 BY SIMILARITY.
FT DISULFID 386 402 BY SIMILARITY.
FT DISULFID 487 551 BY SIMILARITY.
FT DISULFID 515 530 BY SIMILARITY.
FT DISULFID 541 570 BY SIMILARITY.
FT CARBOHYD 106 106 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 116 116 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 174 174 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 182 182 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 267 267 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 514 514 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 556 556 N-LINKED (GLCNAC. .) (POTENTIAL).
FT ACT_SITE 401 401 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 449 449 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 545 545 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT CONFLICT 114 114 N -> K (IN REF. 2).
FT CONFLICT 236 252 NGKHPQEKACNGVND -> MGSTFLRRKPATVMTV
      (IN REF. 1).
SQ SEQUENCE 603 AA; 67234 MW; E2C1D43261C719B8 CRC64;

Query Match      8.4%; Score 76; DB 1; Length 603;
Best Local Similarity 23.6%; Pred. No. 6.5;
Matches 35; Conservative 24; Mismatches 55; Indels 34; Gaps 9;

Qy 17 QEERFPQGL-WTGVMRSCPEQYWDPLLGTCMSCKTI-CNHOSQRTC-----AAFC 66
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 222 QDAEPTSLFQCVNGKHIPQEK-----ACNGVNDGDSDELCKGCGRGNASLC 271
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Qy 67 RSLSCRKQGFYDHLRLDCISACIQGHPKQAYFCENKLRSPVNLPPRLRQRSGEV 126
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 272 KSGVICPDQYKNGEV--DCIT-----GEDESRCF---EDRQQ---NIPKGLARSAQGEA 318
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Qy 127 ENNSDNGRY-QGLEHRSQSEASPALPGL 153
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 319 EIEETEEMLTTPGMDNERKRIKSLPLKL 346
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

RESULT 13
PCK5_RAT
ID PCK5_RAT STANDARD; PRT; 1877 AA.
AC P41413; Q62914;
DT 01-NOV-1995 (Rel. 32, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Proprotein convertase subtilisin/kexin type 5 precursor (EC 3.4.21.-)
DE (Proprotein convertase PC5) (Subtilisin/kexin-like protease PC5)
DE (Convertase PC5) (PC5) (PC6) (Fragments).
GN PCK5.
OS Rattus norvegicus (Rat).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN SEQUENCE FROM N.A. (SHORT ISOFORM), AND TISSUE-SPECIFICITY.
RP TISSUE-Adrenal gland;
RC MEDLINE=93342056; PubMed=8341687;
RX
```

```
RA Lusson J., Vieau D., Hamelin J., Day R., Chretien M., Seidah N.G.;
RT "cDNA structure of the mouse and rat subtilisin/kexin-like PC5: a
RT candidate proprotein convertase expressed in endocrine and
RT nonendocrine cells.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:6691-6695(1993).
RN [2].
RP SEQUENCE OF 1676-1877 FROM N.A. (LONG ISOFORM).
RC TISSUE-Adrenal gland;
RA De Bie I., Marcinkiewicz M., Nakayama K., Lazure C., Seidah N.G.;
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
RN [3].
RP SUBCELLULAR LOCATION.
RX MEDLINE=20214819; PubMed=10749928;
RN Xiang Y., Molloy S.S., Thomas L., Thomas G.;
RT "The PC6B cytoplasmic domain contains two acidic clusters that direct
RT sorting to distinct trans-Golgi network/endosomal compartments.";
RL Mol. Biol. Cell 11:1257-1273(2000).
RN [4].
RP DEVELOPMENTAL EXPRESSION.
RX MEDLINE=97166043; PubMed=9013936;
RN Zheng M., Seidah N.G., Pintar J.E.;
RT "The developmental expression in the rat CNS and peripheral tissues of
RT proteases PC5 and PACE4 mRNAs: comparison with other proprotein
RT processing enzymes.";
RL Dev. Biol. 181:268-283(1997).
CC -1- FUNCTION: LIKELY TO REPRESENT A WIDESPREAD ENDOPEPTIDASE ACTIVITY
CC WITHIN THE CONSTITUTIVE AND REGULATED SECRETORY PATHWAY. CAPABLE
CC OF CLEAVAGE AT THE RX(K/R)R CONSENSUS MOTIF. MAY BE RESPONSIBLE
CC FOR THE MATURATION OF GASTROINTESTINAL PEPTIDES. MAY BE INVOLVED
CC IN THE CELLULAR PROLIFERATION OF ADRENAL CORTEX VIA THE ACTIVATION
CC OF GROWTH FACTORS.
CC -1- CATALYTIC ACTIVITY: RELEASE OF MATURE PROTEINS FROM THEIR
CC PROPOTIENS BY CLEAVAGE OF ARG-XAA-YAA-ARG-1-ZAA BONDS, WHERE XAA
CC CAN BE ANY AMINO ACID AND YAA IS ARG OR LYS.
CC -1- SUBCELLULAR LOCATION: PC5A IS A TYPE I MEMBRANE PROTEIN LOCALIZED TO
CC A PARANUCLEAR POST-GOLGI NETWORK COMPARTMENT IN COMMUNICATION WITH
CC EARLY ENDOSOMES (BY SIMILARITY).
CC -1- ALTERNATIVE PRODUCTS: AT LEAST 2 ISOFORMS: PC5B/LONG (SHOWN HERE)
CC AND PC5A/SHORT; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN THE INTESTINE, BRAIN, ADRENAL
CC GLAND, ANTERIOR PITUITARY, THYROID, OVARIAN, TESTIS AND LUNG.
CC HIGHEST LEVELS ARE FOUND IN THE GUT, DUODENUM, JEJUNUM AND ILEUM.
CC EXPRESSION IS HIGHER IN FEMALE THAN IN MALE REPRODUCTIVE ORGANS.
CC -1- DEVELOPMENTAL STAGE: FIRST DETECTED AT E9 IN HIGHLY RESTRICTED
CC REGIONS OF THE NEURAL TUBE, IN CAUDAL MYOTOMES, AND AT THE
CC MATERNO-EMBRYONIC JUNCTION OF THE UTERUS. AT E10, RESTRICTED
CC EXPRESSION IS DETECTED IN THE OPTIC AND OTIC VESICLES, THE ROOF OF
CC MIDBRAIN, AND TRUNK MYOTOMES. BY MIDGESTATION (E13-E16),
CC EXPRESSION IN THE DEVELOPING NERVOUS SYSTEM HAS EXPANDED TO
CC MULTIPLE REGIONS INCLUDING HIPPOCAMPUS, THALAMUS, HYPOTHALAMUS,
CC BRAIN STEM, AND SPINAL CORD. EXPRESSION IS ALSO DETECTED IN
CC SEVERAL PERIPHERAL ORGAN SYSTEMS, INCLUDING GUT, LUNG, ADRENAL AND
CC KIDNEY PRIMORDIA.
CC -1- DOMAIN: THE PROPEPTIDE DOMAIN ACTS AS AN INTRAMOLECULAR CHAPERONE
CC ASSISTING THE FOLDING OF THE ZMOGEN WITHIN THE ENDOPLASMIC
CC RETICULUM.
CC -1- DOMAIN: AC 1 AND AC 2 (CLUSTERS OF ACIDIC AMINO ACIDS) CONTAIN
CC SORTING INFORMATION. AC 1 DIRECTS TGN LOCALIZATION AND INTERACTS
CC WITH THE TGN SORTING PROTEIN PACS-1.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8; ALSO KNOWN AS THE
CC SUBTILASE FAMILY.
CC -1- SIMILARITY: CONTAINS 1 HOMO B/P DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L14933; AAA99906.1; -.
```

```
DR EMBL: U47014; AAA87888.1; -.
DR PIR: B48225; B48225.
DR HSP: Q99405; 1MPT.
DR MEROPS: S08.076; -.
DR InterPro: IPR002174; Furin-like.
DR InterPro: IPR002884; P_domain.
DR InterPro: IPR000209; Peptidase_s8.
DR Pfam: PF01483; P; 1.
DR Pfam: PF00082; Peptidase_s8; 1.
DR PRINTS: PR00723; SUBTILISIN.
DR ProDom: PD000717; P_domain; 1.
DR SMART: SM00261; FU_6.
DR PROSITE: PS00136; SUBTILASE_ASP; 1.
DR PROSITE: PS00137; SUBTILASE_HIS; 1.
DR PROSITE: PS00138; SUBTILASE_SER; 1.
DR Hydrolase: Serine protease; Glycoprotein; Zymogen; Signal;
KW Cleavage on pair of basic residues; Alternative splicing; Repeat;
KW Transmembrane.
FT SIGNAL 1 34 BY SIMILARITY.
FT PROPEP 35 116 BY SIMILARITY.
FT CHAIN 117 1877 PROPROTEIN CONVERTASE SUBTILISIN/KEXIN
FT TYPE 5.
FT DOMAIN 117 1768 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1769 1789 POTENTIAL.
FT DOMAIN 1790 1877 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 117 452 CATALYTIC.
FT DOMAIN 464 602 HOMO B.
FT DOMAIN 638 1753 CYS-RICH MOTIF (CRM) REGION.
FT DOMAIN 1825 1844 AC 1.
FT DOMAIN 1856 1877 AC 2.
FT SITE 116 117 CLEAVAGE (AUTO-) (BY SIMILARITY).
FT SITE 521 523 CELL ATTACHMENT SITE (POTENTIAL).
FT ACT_SITE 173 173 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 214 214 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 388 388 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT CARBOHYD 227 227 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 383 383 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 667 667 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 754 754 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 804 804 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 854 854 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1710 1710 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1732 1732 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPPLIC 878 915 XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX ->
FT ATEESWAEGGFCMLVKKNLCORVKVQLQCCCKTCTFQG
FT (IN ISOFORM PC5A).
FT VARSPPLIC 916 1877 MISSING (IN ISOFORM PC5A).
FT SEQUENCE 1877 AA; 207888 MW; 890955DC60534444 CRC64;

Query Match 8.4%; Score 76; DB 1; Length 1877;
Best Local Similarity 25.0%; Pred. No. 20;
Matches 20; Conservative 5; Mismatches 35; Indels 20; Gaps 3;

QY 27 TGVAMRSCPEEQYWDPLLGTGCMCKTICNHQSORTCAAFCSRISCRK----- 73
DB 719 TSSCAQCPGSGYODIKKNTGCKSENC-----KTCTGFHNCTECKGLSLOGSRCSVTC 773
QY 74 EQGKFYDHLRLDCISASC 93
DB 774 EDGQFFSG--HDCQPCRHC 791

RESULT 14
YKT5_YEAST
ID YKT5_YEAST STANDARD; PRT; 427 AA.
AC P36046;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE Hypothetical 47.4 kDa protein in PAS1-MST1 intergenic region.
GN YKL195W.
OS Saccharomyces cerevisiae (Baker's yeast).
```

```
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA Mala e Silva A., Bossier P., Villela C., Fernandes L., Soares H.,
RA Guerreiro P., Rodrigues-Pousada C.;
RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z28195; CAA82039.1; -.
DR PIR: S38032; S38032.
DR SGD: S0001678; YKL195W.
KW Hypothetical protein.
SQ SEQUENCE 427 AA; 47416 MW; 44386D250DE5DED4 CRC64;

Query Match 8.3%; Score 75.5; DB 1; Length 427;
Best Local Similarity 26.9%; Pred. No. 5.2;
Matches 32; Conservative 20; Mismatches 52; Indels 15; Gaps 5;

QY 53 ICHQSORTCAA---FCRSLSRCRKEQGFYDHLRLDCISCA-SICGQHPKQCAFCENKL 108
DB 47 LCRHQRRLMASSPQFGRRNSQKRTAG-----FMGILSMAGALYFAPNRKPLFASRKV 101
QY 109 RSPVNLPPPEL-----RRQRSGEVENN--SDNSGRYQGLEHGRGSEASPALPGLKLSADOVA 161
DB 102 ESKTAEELSSGGEQSPENEDNNNSKSDENGDDNDKNDTEAGPQLGDKKICASKVA 160

RESULT 15
LMB_HIRME
ID LMB_HIRME STANDARD; PRT; 400 AA.
AC Q25092;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Laminin B-chain (Fragment).
OS Hirudo medicinalis (Medicinal leech).
OC Eukaryota; Metazoa; Annelida; Clitellata; Hirudinida; Hirudinea;
OC Arynchobdellida; Hirudiniformes; Hirudinidae; Hirudo.
OX NCBI_TaxID=6421;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95370770; PubMed=7643070;
RA Luecke A.E., Dickerson I.M., Muller K.J.;
RT "In situ hybridization reveals transient laminin B-chain expression
RT by individual glial and muscle cells in embryonic leech central
RT nervous system.";
RL J. Neurobiol. 27:1-14(1995).
CC -!- FUNCTION: BINDING TO CELLS VIA A HIGH AFFINITY RECEPTOR, LAMININ
CC IS THOUGHT TO MEDIATE THE ATTACHMENT, MIGRATION & ORGANIZATION OF
CC CELLS INTO TISSUES DURING EMBRYONIC DEVELOPMENT BY INTERACTING
CC WITH OTHER EXTRACELLULAR MATRIX COMPONENTS.
CC -!- SUBUNIT: LAMININ IS A COMPLEX GLYCOPROTEIN, CONSISTING OF THREE
CC DIFFERENT POLYPEPTIDE CHAINS (ALPHA, BETA, GAMMA), WHICH ARE BOUND
CC TO EACH OTHER BY DISULFIDE BONDS INTO A CROSS-SHAPED MOLECULE
CC COMPRISING ONE LONG & THREE SHORT ARMS WITH GLOBULES AT EACH END.
CC -!- TISSUE SPECIFICITY: INDIVIDUAL GLIAL AND MUSCLE CELLS.
CC -!- DEVELOPMENTAL STAGE: EMBRYONIC DEVELOPMENT.
CC -!- DOMAIN: THE ALPHA-HELICAL DOMAINS I AND II ARE THOUGHT TO INTERACT
CC WITH OTHER LAMININ CHAINS TO FORM A COILED COIL STRUCTURE.
CC -!- SIMILARITY: CONTAINS AT LEAST 3 LAMININ EGF-LIKE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; U34921; AAC46862.1; .
DR HSSP; P02468; IKLO.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR002049; Laminin_EGF.
DR Pfam; PF00053; Laminin_EGF; 2.
DR SMART; SM00180; EGF_Lam; 2.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01248; LAMININ_TYPE_EGF; 2.
KW Glycoprotein; Extracellular matrix; Repeat; Laminin EGF-like domain;
KW Coiled coil.
FT NON_TER 1 1
FT DOMAIN <1 5 LAMININ EGF-LIKE.
FT DOMAIN 6 53 LAMININ EGF-LIKE.
FT DOMAIN 54 100 LAMININ EGF-LIKE.
FT DOMAIN 101 >400 DOMAIN II AND I.
FT DOMAIN 140 235 COILED COIL (POTENTIAL).
FT DOMAIN 353 >400 COILED COIL (POTENTIAL).
FT DISULFID 6 18 BY SIMILARITY.
FT DISULFID 8 25 BY SIMILARITY.
FT DISULFID 27 36 BY SIMILARITY.
FT DISULFID 39 51 BY SIMILARITY.
FT DISULFID 54 66 BY SIMILARITY.
FT DISULFID 56 73 BY SIMILARITY.
FT DISULFID 75 84 BY SIMILARITY.
FT DISULFID 87 98 BY SIMILARITY.
FT DISULFID 101 101 INTERCHAIN (PROBABLE).
FT DISULFID 104 104 INTERCHAIN (PROBABLE).
FT CARBOHYD 160 160 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 175 175 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 216 216 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 266 266 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 283 283 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 310 310 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 356 356 N-LINKED (GLCNAC. .) (POTENTIAL).
FT NON_TER 400 400
SQ SEQUENCE 400 AA; 43262 MW; 4330ABE05A61B012 CRC64;

Query Match 8.3%; Score 75; DB 1; Length 400;
Best Local Similarity 23.3%; Pred. No. 5.4;
Matches 41; Conservative 17; Mismatches 66; Indels 52; Gaps 10;
Qy 9 RGRSRVDQERFPQGLWTGYAMRSCPEEQYWDPLLTGTCMCKTICNHQSORT----- 61
Db 31 RGRDR-----CSQCPMSWMDPFLG-CKSC--TCNPDGARSLYCNKVT 70
Qy 62 ----CAAFCSRSLSC-RKEQGFYDHLRLDCISASICGQHPKQAYF-----CENKLRSP 111
Db 71 GQCECPRGVTGLNCDRCRGTY--GALPQCIPGCECFDNDKLIQLRDEAAQURIGTE 128
Qy 112 VNL--PPELRQRSGEVEN-----NSDN--SGRYQGLEHRGSEASPALPGLK 154
Db 129 IKLSGPPGFAFAKFEELBQVLMDKMSHVNSANVSSDQLENIDQELDNLSSKLDK 184

Search completed: June 25, 2002, 16:22:44
Job time: 629 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 25, 2002, 16:24:27 ; Search time 89.98 Seconds
(without alignments)
319.150 Million cell updates/sec

Title: US-09-854-864-15
Perfect score: 909
Sequence: 1 MSGIGRRGRRSRVDQER.....SPALPGLKLSADQVALVYST 166

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 17299429 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_19.*
1: sp-archaea.*
2: sp-bacteria.*
3: sp-fungi.*
4: sp-human.*
5: sp-invertebrate.*
6: sp-mammal.*
7: sp-mhc.*
8: sp-organelle.*
9: sp-phage.*
10: sp-plant.*
11: sp-rodent.*
12: sp-virus.*
13: sp-vertebrate.*
14: sp-unclassified.*
15: sp-rvrius.*
16: sp-bacteriap.*
17: sp-archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	909	100.0	293	4	O14836
2	419.5	46.1	249	11	Q9DBZ3
3	419.5	46.1	249	11	Q9ET35
4	140	15.4	156	6	Q9N146
5	93	10.2	840	4	Q96H26
6	93	10.2	1240	4	O15030
7	91	10.0	1704	5	Q94446
8	89	9.8	839	5	O10652
9	89	9.8	1878	4	Q9UIF9
10	88.5	9.7	565	5	Q9W241
11	87	9.6	272	4	Q9BX14
12	87	9.6	292	4	Q96K87
13	86.5	9.5	310	10	Q03108
14	86.5	9.5	353	10	Q03107
15	86	9.5	297	4	Q9HAV5
16	86	9.5	1299	5	Q26489

ALIGNMENTS

RESULT 1

```
O14836          267 13 Q9I9M0
17 85 9.4 1137 4 Q9H8C1
18 85 9.4 1137 4 Q9H8C1
19 85 9.4 1918 4 Q9BOM7
20 85 9.4 1925 4 Q9P2E3
21 84.5 9.3 702 5 Q9VH96
22 83 9.1 820 10 Q9FFK8
23 83 9.1 971 11 Q920J1
24 83 9.1 13288 6 O18758
25 82.5 9.1 913 4 Q96EP4
26 82.5 9.1 1362 13 Q9PVZ4
27 82 9.0 239 13 Q9OZL0
28 82 9.0 868 5 Q9YIV3
29 82 9.0 932 11 Q62030
30 82 9.0 1650 11 Q9QVT6
31 81.5 9.0 316 11 Q922H9
32 81 8.9 704 3 Q74567
33 81 8.9 971 4 Q95980
34 81 8.9 1140 4 Q96KG7
35 80 8.8 108 16 O07571
36 80 8.8 330 5 O18118
37 80 8.8 996 11 Q924X6
38 80 8.8 1664 5 Q9TVQ2
39 80 8.8 1792 13 O57484
40 79.5 8.7 98 5 Q16939
41 79.5 8.7 593 10 Q9SEW4
42 79.5 8.7 1963 5 Q9VY56
43 79.5 8.7 2189 5 Q9BI05
44 79 8.7 217 11 Q9CSB2
45 79 8.7 269 17 O29751
```

PRELIMINARY; PRT; 293 AA.

AC O14836; PRELIMINARY; PRT; 293 AA.

DT O1-JAN-1998 (TREMBLrel. 05, Created)

DT O1-JAN-1998 (TREMBLrel. 05, Last sequence update)

DT O1-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE TRANSMEMBRANE ACTIVATOR AND CAML INTERACTOR.

GN TACI.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=97458245; PubMed=9311921;

RA von Bulow G.U., Bram R.J.;

RT "NF-AT activation induced by a CAML-interacting member of the tumor necrosis factor receptor superfamily.";

RL Science 278:138-141(1997).

DR EMBL; AF023614; AAC51790.1; -.

DR InterPro; IPR001368; TNFR_C6.

DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.

SQ SEQUENCE 293 AA; 31816 MW; 411799F3DE17A5EB CRC64;

Query Match 100.0%; Score 909; DB 4; Length 293;
Best Local Similarity 100.0%; Pred. No. 1.2e-94;
Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MSGIGRRGRRSRVDQERFPQGLWTGTVAMRSCPEEQYWDPLLGTCMSCKTICNHQSQR 60
   |||||
Db 1 MSGIGRRGRRSRVDQERFPQGLWTGTVAMRSCPEEQYWDPLLGTCMSCKTICNHQSQR 60
   |||||
QY 61 TCAAFRCSLSCREQGFYDHLRLDCISCSICGQHPKQCAIFCNKLRSPVNLPPELRR 120
   |||||
Db 61 TCAAFRCSLSCREQGFYDHLRLDCISCSICGQHPKQCAIFCNKLRSPVNLPPELRR 120
   |||||
QY 121 QRSGEVNNSDNNGRYOGLHRCGSEASPALPGLKLSADQVALVYST 166
```


(FRAGMENT).
DE Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LYMPH, AND LYMPHOMA;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC008965; AAH08965.1;
FT NON_TER 1
SQ SEQUENCE 840 AA; 3F8147D9B7034B45 CRC64;

Query Match 10.2%; Score 93; DB 4; Length 840;
Best Local Similarity 25.8%; Pred. No. 0.075;
Matches 40; Conservative 17; Mismatches 68; Indels 30; Gaps 7;

QY 22 PQGLWTCVAMRSCPEEQYWDPLLTGTCMSCKTI--CNHQSQRTCA-----AFCRSLSCR 72
Db 561 PEGTTTTEISYEITPRIRVWRQTLERCRAAQVCLCLGLQLERSIAWEKSVNKKVTC--LVCR 618
QY 73 KEQKGFYDHLRLDCISCASICG---QHPKQCA-----YFC-----ENKLRSPVNLPP 116
Db 619 KGDN---DEFLLCDGCDRGCHYCHRPKMEAVPEGDMFCTVCLAAQQVEGFTQKPGFPK 675
QY 117 ELRRQRSGEVENSDNSGRYOGLEHGRSEASPALP 151
Db 676 RGQKRKSGYSLNFGDGRRRVLLRGRESPAAGP 710

RESULT 6
OI15030
ID OI15030 PRELIMINARY; PRT; 1240 AA.
AC OI15030;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE KIAA0314 PROTEIN (FRAGMENT).
GN KIAA0314.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE=BRIN.
RC MEDLINE=97349984; PubMed=9205841;
RA Nagase T., Ishikawa K., Nakajima D., Ohira M., Seki N., Miyajima N.,
Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. VII.
The complete sequences of 100 new cDNA clones from brain which can
code for large proteins in vitro";
RL DNA Res. 4:141-150(1997).
DR EMBL: AB002312; BAA20773.1; -;
DR InterPro: IPR000637; AT_hook.
DR InterPro: IPR001487; Bromodomain.
DR InterPro: IPR004022; DDT.
DR InterPro: IPR001965; PHD.
DR Pfam: PF02178; AT_hook; 3.
DR Pfam: PF00439; bromodomain; 1.
DR Pfam: PF02791; DDT; 1.
DR Pfam: PF00628; PHD; 1.
DR PRINTS: PR00929; ATHOOK.
DR PRINTS: PR00503; BROMODOMAIN.
DR SMART: SM00384; AT_hook; 3.
DR SMART: SM00297; BROMO; 1.
DR PROSITE: PS50014; BROMODOMAIN_2; 1.
FT NON_TER 1
SQ SEQUENCE 1240 AA; 140417 MW; 20BDBFF1AA6BC5A1 CRC64;

Query Match 10.2%; Score 93; DB 4; Length 1240;
Best Local Similarity 25.8%; Pred. No. 0.12;
Matches 40; Conservative 17; Mismatches 68; Indels 30; Gaps 7;

QY 22 PQGLWTCVAMRSCPEEQYWDPLLTGTCMSCKTI--CNHQSQRTCA-----AFCRSLSCR 72
Db 961 PEGTTTTEISYEITPRIRVWRQTLERCRAAQVCLCLGLQLERSIAWEKSVNKKVTC--LVCR 1018
QY 73 KEQKGFYDHLRLDCISCASICG---QHPKQCA-----YFC-----ENKLRSPVNLPP 116
Db 1019 KGDN---DEFLLCDGCDRGCHYCHRPKMEAVPEGDMFCTVCLAAQQVEGFTQKPGFPK 1075
QY 117 ELRRQRSGEVENSDNSGRYOGLEHGRSEASPALP 151
Db 1076 RGQKRKSGYSLNFGDGRRRVLLRGRESPAAGP 1110

RESULT 7
Q94446
ID Q94446 PRELIMINARY; PRT; 1704 AA.
AC Q94446;
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE 220 KDA SILK PROTEIN.
GN SP220.
OS Chironomus thummi (midge).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera;
OC Chironomidae; Chironominae; Chironominae; Chironomus.
OX NCBI_TaxID=7154;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SALIVARY GLAND;
RA Case S.T., Cox C., Bell W.C., Hoffman R.T., Martin J., Hamilton R.;
RT "Extraordinary conservation of cysteines among homologous Chironomus
silk proteins sp185 and sp220";
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: U54641; AAA99804.1; -;
DR InterPro: IPR004153; CXCX.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR000853; Mllthion_nemat.
DR Pfam: PF03128; CXCX; 69.
DR PRINTS: PR00876; MTNEMATODE.
DR PROSITE: PS00022; EGF_1; UNKNOWN_2.
DR PROSITE: PS01186; EGF_2; UNKNOWN_1.
SQ SEQUENCE 1704 AA; 185746 MW; 3A3F20247C8F1E28 CRC64;

Query Match 10.0%; Score 91; DB 5; Length 1704;
Best Local Similarity 21.4%; Pred. No. 0.28;
Matches 31; Conservative 22; Mismatches 44; Indels 48; Gaps 7;

QY 32 RSCPEEQYWD-----PLLGTG-----MSCRTTCNHQ-----SQTCAAFC 66
Db 1357 QTCPAGOSWDSQTCQSCPATGCTGAQFWCAKQKCVQENCKGPKVFDQTSQSCQC 1416
QY 67 -RSLSCRKEQ---GKFYDH-----LLRDCISCASICGQHPKQC---AYF 103
Db 1417 PKNMQPPGECTAGRTWDDATCTEKCATVPNCESPMVFDQATCGCKGNPKLPADKVW 1476
QY 104 CENKLRSPVNLPPELRRQRSGEVEN 128
Db 1477 CDKCKQAVCSLPPTITQCPYSGQTYN 1501

RESULT 8
Q10652
ID Q10652 PRELIMINARY; PRT; 839 AA.
AC Q10652;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)

DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE ZINC FINGER PROTEIN CEZF.
 GN CEZF.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 CC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=BRISTOL N2;
 RC MEDLINE=96003854; PubMed=7568208;
 RX Saha V., Chaplin T., Gregorini A., Ayton P., Young B.D.;
 RA "The leukemia-associated-protein (LAP) domain, a cysteine-rich motif,
 RT is present in a wide range of proteins, including MLL, AF10, and MLL76
 RT proteins.";
 RL Proc. Natl. Acad. Sci. U.S.A. 92:9737-9741(1995).
 CC -1- SIMILARITY: CONTAINS A CYS-RICH (PHD-FINGER) DOMAIN SIMILAR TO
 CC THOSE FROM HAT3.1, MAIZE HOXA AND PARSLEY PRH.
 DR EMBL; U20555; AAC46918.1; -;
 DR InterPro; IPR001965; PHD.
 DR SMART; PF00628; PHD; 1.
 DR SMART; SM00249; PHD; 2.
 KW DNA-binding; zinc-finger; Metal-binding.
 FT DOMAIN 8 54 CYS-RICH (PHD-FINGER).
 FT DOMAIN 619 655 LEUCINE ZIPPER.
 FT ZN_FING 127 151 POTENTIAL.
 SQ SEQUENCE 839 AA; 89439 MW; AFBFFB9D1D35B4B8 CRC64;

Query Match 9.8%; Score 89; DB 5; Length 839;
 Best Local Similarity 32.5%; Pred. No. 0.21; Indels 20; Gaps 5;
 Matches 27; Conservative 8; Mismatches 28; Indels 20; Gaps 5;

Qy 34 CPEQYWDPLGLTGMSC-KTICNHQSORTCAAFRCRSLSRCREQGKFYDHLRLDCISCASI 92
 Db 130 CNEERPNDKAGCMSCNKSCKRSFHYTCAQ-RKGLLC---EEGAI----- 172
 Qy 93 CGQHPKQCAFCENKLRSPVNL 115
 Db 173 -SRNVKCYGY-CENHLKAINDP 193

RESULT 9
 Q9UIF9 PRELIMINARY; PRT; 1878 AA.
 AC Q9UIF9;
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE BROMODOMAIN ADJACENT TO ZINC FINGER DOMAIN 2A.
 GN BAZ2A.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=20130112; PubMed=10662543;
 RX Jones M.H., Hamana N., Nezu J., Shimane M.;
 RA "A novel family of bromodomain genes.";
 RL Genomics 63:40-45(2000).
 DR EMBL; AB032254; BAA89211.1; -;
 DR InterPro; IPR000637; AT_hook.
 DR InterPro; IPR001487; Bromodomain.
 DR InterPro; IPR004022; DDT.
 DR InterPro; IPR001739; MBD.
 DR InterPro; IPR001965; PHD.
 DR Pfam; PF02178; AT_hook; 4.
 DR Pfam; PF00439; bromodomain; 1.
 DR Pfam; PF02791; DDT; 1.
 DR Pfam; PF01429; MBD; 1.
 DR Pfam; PF00628; PHD; 1.
 DR PRINTS; PR00503; BROMODOMAIN.

DR SMART; SM00384; AT_hook; 3.
 DR SMART; SM00297; BROMO; 1.
 DR SMART; SM00391; MBD; 1.
 DR SMART; SM00249; PHD; 1.
 DR PROSITE; PS0014; BROMODOMAIN_2; 1.
 SQ SEQUENCE 1878 AA; 208639 MW; 12683AFE636A93A6 CRC64;

Query Match 9.8%; Score 89; DB 4; Length 1878;
 Best Local Similarity 23.3%; Pred. No. 0.53;
 Matches 37; Conservative 18; Mismatches 56; Indels 38; Gaps 6;

Qy 22 PQGLWTGVAMRSPPEQYWDPLGLTGMSCKTIC-----NHQSORTCAAFRCRS 68
 Db 1599 PEGTTTSEYITPRIRIWRQTLQCRSAHVCLGLHLERSIAWKSIVNKVC----- 1652
 Qy 69 LSCRKEQGFYDHLRLDCISCASICG---QHPKQCA-----YFC-----ENKLRSPV 112
 Db 1653 LVCRRKGN---DEFLLLCDGCDRGCHYCHRPKMEAVPEGDMFCTVCLAQQVEGEFTQKP 1709
 Qy 113 NLPPELRRQRSGEVNNSDNGRYQGLEHGRGSEASPALP 151
 Db 1710 GFPRGQKRSYSINFGDGRRRVLLKGRSPAGP 1748

RESULT 10
 Q9W241 PRELIMINARY; PRT; 665 AA.
 AC Q9W241;
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-OCT-2001 (Tremblrel. 18, Last annotation update)
 DE CG12489 PROTEIN.
 GN CG12489.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 CC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 CC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=BERKELEY.
 RC MEDLINE=20196006; PubMed=10731132;
 RX Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Vandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottiler P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,

RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "the genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 DR EMBL; AE003457; AAF46858.1;
 DR FlyBase; FBgn0034738; CG12489.
 DR InterPro; IPR000299; Band 4.1.
 DR InterPro; IPR001841; Znf_Ting.
 DR Pfam; PF00373; Band_4.1.
 DR SMART; SM00295; B41; 1.
 DR SMART; SM00184; RING; 1.
 DR PROSITE; PS50057; BAND_41.3; 1.
 SQ SEQUENCE 665 AA; 71851 MW; D0281FBA5BA45683 CRC64;

Query Match 9.7%; Score 88.5; DB 5; Length 665;
 Best Local Similarity 25.9%; Pred. No. 0.18;
 Matches 43; Conservative 19; Mismatches 65; Indels 39; Gaps 9;
 QY 2 SGLGRSRGSRVQDQERFPQGLWTGVAMRSCPEEQYWDPLLTGTCMCKTICNHQSQRT 61
 Db 499 TGAGGSMAG---KIDLAIREKEA-----REAAIERCVDTRISEAMOCK-ICMDRAINT 547
 QY 62 CAAFCSRSLSCRKEGKGFYDHLRLDCISCATCGQHPKQCAFCKNKLSPVN--LPPELR 119
 Db 548 VFNFC-----C-----HWIA-CAQCAARCSNCPN-----CRVKTSVVKYVLPPELR 588
 QY 120 QRGRGEVNNDSNG-----RVQGLEHRGSEASPALGPLKLSAD 158
 Db 589 TSOTGSGATTSSSSIMGDGOVEQLQLQQLDEISAAPASLEAGAD 634

RESULT 11
 ID Q9BX4 PRELIMINARY; PRT; 272 AA.
 AC Q9BX4;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
 DE THROMBOSPONDIN.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Mao Y., Xie Y., Zhou Z., Zhao W., Zhao S., Wang W., Huang Y., Wang S.,
 RA Tang R., Chen X., Wu C.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF251057; AAK34947.1;
 DR InterPro; IPR00174; Furin-like.
 DR InterPro; IPR000884; TSPI.
 DR Pfam; PF00090; tsp_1; 1.
 DR SMART; SM00361; FU; 2.
 DR SMART; SM00209; TSPI; 1.
 DR PROSITE; PS50092; TSPI; 1.
 SQ SEQUENCE 272 AA; 30928 MW; CACAE6B7E781189 CRC64;

Query Match 9.6%; Score 87; DB 4; Length 272;
 Best Local Similarity 24.6%; Pred. No. 0.098;
 Matches 42; Conservative 18; Mismatches 59; Indels 52; Gaps 11;
 QY 28 GVAMRSCPEQYWD--PLLGTMCCKTICNHQSQRTCAAFCSRSLSCRKEGKFYDHL--- 82
 Db 74 GVCILSSCPGSGYGYTRYPDKCTCKACDC-----TC--FNKNF-CTKCKSGFYVHLGKC 125
 QY 83 LRDC-----ISCASI-----CGQHPKQCAF--FCENKLRSVP----- 112

Db 126 LDNCPGEGLEANNHTMECVSIHVCEVSEWNPWSPTCKGKTCGFRGTETRVREIIQHPSA 185
 QY 113 --NLPPELRRQRSGEVENNSDNGR--YQGLEHR-----GSEASPALPLK 154
 Db 186 KGNLCPTNETRKTCTVQRKKCKQGERGKKGRERKKPKNKGESKEAIPDSK 236
 RESULT 12
 ID Q96K87 PRELIMINARY; PRT; 292 AA.
 AC Q96K87;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE CDNA FLJ14440 FTS, CLONE HEMBB1000915, WEAKLY SIMILAR TO
 DE SUBTILISIN-LIKE PROTEASE PACE4 PRECURSOR (EC 3.4.21.-).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=EMBRVO;
 RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
 RA Nishikawa T., Nagai K., Sugano S., Takahashi-Fujii A., Hara H.,
 RA Tanase T., Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K.,
 RA Arita M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J.,
 RA Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y., Oshima A.;
 RT "NEDO human cDNA sequencing project.";
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK027346; BAB55051.1;
 SQ SEQUENCE 292 AA; 33243 MW; 01E2774AC3D4A6F8 CRC64;

Query Match 9.6%; Score 87; DB 4; Length 292;
 Best Local Similarity 24.6%; Pred. No. 0.11;
 Matches 42; Conservative 18; Mismatches 59; Indels 52; Gaps 11;
 QY 28 GVAMRSCPEQYWD--PLLGTMCCKTICNHQSQRTCAAFCSRSLSCRKEGKFYDHL--- 82
 Db 74 GVCILSSCPGSGYGYTRYPDKCTCKACDC-----TC--FNKNF-CTKCKSGFYVHLGKC 125
 QY 83 LRDC-----ISCASI-----CGQHPKQCAF--FCENKLRSVP----- 112
 Db 126 LDNCPGEGLEANNHTMECVSIHVCEVSEWNPWSPTCKGKTCGFRGTETRVREIIQHPSA 185
 QY 113 --NLPPELRRQRSGEVENNSDNGR--YQGLEHR-----GSEASPALPLK 154
 Db 186 KGNLCPTNETRKTCTVQRKKCKQGERGKKGRERKKPKNKGESKEAIPDSK 236

RESULT 13
 ID Q03108 PRELIMINARY; PRT; 310 AA.
 AC Q03108;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE CATHEPSIN B (FRAGMENT).
 OS Triticum aestivum (Wheat).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
 OC Triticeae; Triticum.
 OX NCBI_TaxID=4565;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VAR. CHINESE SPRING;
 RX MEDLINE=93258430; PubMed=1302642;
 RA Cejudo F.J., Murphy G., Chino C., Baulcombe D.C.;
 RT "A gibberellin-regulated gene from wheat with sequence homology to
 RT cathepsin B of mammalian cells.";
 RL Plant J. 2:937-948(1992).
 DR EMBL; X66014; CAA46812.1;

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 25, 2002, 16:16:08 ; Search time 88.08 Seconds
(without alignments)
209.335 Million cell updates/sec

Title: US-09-854-864-15
Perfect score: 909
Sequence: 1 MSGLSRRGGRSRVDQER.....SPALPLKLSADQVALVYST 166

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues
Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802:*
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	909	100.0	166	19 AAW75785	Human lymphocyte s
2	909	100.0	265	22 AAE09244	Human TACI splice
3	909	100.0	293	19 AAW75783	Human lymphocyte s
4	909	100.0	233	21 AAB36312	Human neutrokinine-a
5	909	100.0	293	21 AAY94000	A transmembrane ac
6	909	100.0	293	22 AAE09240	Human TACI protein
7	909	100.0	293	22 AAY19114	Human tumour necro
8	599	65.9	247	21 AAY93998	Human BR43x2, an i
9	419.5	46.1	249	21 AAY94006	A murine tnfa, a
10	95	10.5	197	21 AAB01421	Human TANGO 140-2.
11	93	10.2	1589	22 AAM42025	Human polypeptide

12	93	10.2	1727	22 AAB95554	Human protein sequ
13	93	10.2	1878	22 AAM40239	Human polypeptide
14	92	10.1	266	22 AAM39716	Human polypeptide
15	89	9.8	1878	19 AAW81170	Human BAZ2-alpha p
16	88.5	9.7	665	22 ABE68244	Drosophila melanog
17	88	9.7	795	22 ABG07025	Novel human diagno
18	87.5	9.6	463	20 AAY13392	Amino acid sequenc
19	87.5	9.6	463	21 AAB01373	Neuron-associated
20	87.5	9.6	463	21 AAY95343	Human PRO328 antit
21	87.5	9.6	463	22 AU12351	Human PRO328 poly
22	87.5	9.6	463	22 AAB88408	Human membrane or
23	87.5	9.6	463	22 AAB80260	Human PRO328 prote
24	87.5	9.6	463	22 AAB53088	Human angiogenesis
25	87	9.6	251	22 AAE13153	Human mature stem
26	87	9.6	265	22 AAE13163	Human secreted pro
27	87	9.6	272	22 AAE13150	Human stem cell gr
28	87	9.6	272	22 AAE13168	Human stem cell gr
29	87	9.6	272	22 AAM78328	Human protein SEQ
30	87	9.6	272	22 AAB99220	Human thrombospond
31	87	9.6	273	22 AAE13151	Human stem cell gr
32	87	9.6	292	20 AAW85607	Secreted protein c
33	87	9.6	292	22 AAE13170	Human SCR-1 relate
34	87	9.6	292	22 AAB93875	Human protein sequ
35	86	9.5	173	22 AAU03118	Composite protein
36	86	9.5	206	21 AAB01420	Human TANGO 140-1.
37	86	9.5	231	21 AAY77468	Human Rank-like pr
38	86	9.5	267	22 AAU03114	Human uterine myom
39	86	9.5	269	22 AAU03106	Human uterine myom
40	86	9.5	297	22 AAU03113	Human uterine myom
41	86	9.5	297	22 AAB29534	Human TNFR homolog
42	86	9.5	299	21 AAB30547	Amino acid sequenc
43	86	9.5	299	21 AAB33477	Human PRO5727 prot
44	86	9.5	299	22 AAU03116	Composite protein
45	86	9.5	299	22 AAB29533	Human TNFR homolog

ALIGNMENTS

RESULT 1	
AAW75785	
ID AAW75785 standard; Protein; 166 AA.	
XX AC	
XX AAW75785;	
DT 18-JAN-1999 (first entry)	
XX Human lymphocyte surface receptor extracellular domain.	
DE TACI; transmembrane activator and CAML-interactor;	
XX calcium signal-modulating cyclophilin ligand; human;	
KW lymphocyte surface receptor; human; B-cell; B lymphocyte;	
KW infection; cancer; rheumatoid arthritis; autoimmune disease;	
KW glomerulonephritis; immunosuppressive; graft versus host disease;	
KW transplant rejection; therapy; signal transduction	
XX Homo sapiens.	
OS	
XX	
PN WO9839361-A1.	
XX	
PD 11-SEP-1998.	
XX	
PF 03-MAR-1998; 98WO-US04270.	
XX	
PR 03-MAR-1997; 97US-0810572.	
XX	
PA (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.	
XX	
PI Bram RJ, Von Bulow G;	
XX	
DR WPI; 1998-506346/43.	
DR N-PSDB; AAW57330.	
XX	

PT New isolated transmembrane activator protein - used to develop
 PT products for treating e.g. infections, cancers, autoimmune and
 PT inflammatory conditions, transplant rejection or graft-versus-host
 PT disease
 XX
 PS Claim 8; Page 73; 89pp; English.
 XX
 CC This is the amino acid sequence of the N-terminal, i.e. the
 CC extracellular, domain of novel human transmembrane activator and
 CC CAML-interactor (TACI) protein (see AAW75783). TACI is a lymphocyte
 CC receptor protein that is involved in the calcium activation pathway.
 CC It is normally present in B-lymphocytes, and to a much lesser extent
 CC in immature T-lymphocytes, and can therefore be targeted to
 CC specifically regulate B cell responses without affecting T cell
 CC activity. The extracellular domain of TACI functions as a binding
 CC site for a ligand that stimulates the activation of the cell by
 CC inducing the binding of the C-terminal portion (see AAW75784) of
 CC TACI to the N-terminal domain of CAML. A recombinant form of the
 CC extracellular portion of TACI acts as a dominant-negative or
 CC blocking agent and acts to suppress the immune system. It can be
 CC used to treat or prevent autoimmune disease, graft rejection or
 CC graft versus host disease. The extracellular region is also used
 CC in a claimed method for identifying a ligand for TACI, in which
 CC binding of a candidate molecule is determined by detecting cellular
 CC activation of the AP-1, CAMP or NF- κ B pathway, of NF-AT
 CC transcription factor, or of NF-AT dependent transcription.
 XX
 SQ Sequence 166 AA;

Query Match 100.0%; Score 909; DB 19; Length 166;
 Best Local Similarity 100.0%; Pred. No. 2.3e-83;
 Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MSGLGRSRGRSRVDOERFPQGLWTGVAMRSCPEEQYWPDLTGTCMSCKTICNHQSQR 60
 DB 1 msglgrsrgrsrvdqeerfpqglwtgvamrscpeeqywdplgtcmstctcnhqsr 60
 QY 61 TCAAFCSLSCKRKGKGYDHLRLDCISCSICGQHPKQCAFCENKURSPVNPPELRR 120
 DB 61 tcaafcslsckrkggkfydhllrdciscaslcgghpkqcafcenklsrsvnpnlpeirr 120
 QY 121 QRSGEVENNSDNGRYQGLEHRGSEASPALPGLKLSADQVALVYST 166
 DB 121 qrsgevennsdngryqglehrgseaspalp9lklksadqvalvyst 166

RESULT 2
 AAE09244
 ID AAE09244 standard; Protein; 265 AA.
 XX
 AC AAE09244;
 XX
 DT 19-NOV-2001 (first entry)
 XX
 DE Human TACI splice variant protein.
 XX
 KW Human; TNF; tumour necrosis factor; TALL-1; APRIL; TNF receptor;
 KW TNFR; TACI; BCMA; therapy; cancer; leukaemia; myeloma; lymphoma;
 KW autoimmune disease; rheumatoid arthritis; multiple sclerosis;
 KW psoriasis.
 XX
 OS Homo sapiens.
 XX
 PN WO200160397-A1.
 XX
 PD 23-AUG-2001.
 XX
 PF 28-NOV-2000; 2000WO-US232378.
 XX
 PR 16-FEB-2000; 2000US-0182938.
 PR 22-AUG-2000; 2000US-0226986.
 XX

PA (GETH) GENENTECH INC.
 XX
 PI Ashkenazi AJ, Dodge KH, Grewal I, Kim KJ, Marsters SA, Pitti RM;
 PI Yan M;
 XX
 DR WPI; 2001-541628/60.
 XX
 CC Inhibiting or neutralizing TALL-1 or APRIL polypeptide biological
 CC activity, for treating autoimmune disorders and cancer, comprises
 CC exposing the cells to TALL-1 or APRIL polypeptide agonists or
 CC antagonists
 XX
 PS Example 1; Fig 6; 160pp; English.
 XX
 CC The invention relates to methods of using one or more agonists or
 CC antagonists to modulate the activity of the members of TNF (tumour
 CC necrosis factor) especially TALL-1, APRIL and TNF receptor (TNFR)
 CC e.g. TACI or BCMA. The method is useful for treating pathological
 CC conditions or diseases associated with increased TALL-1 and APRIL
 CC expression or activity. TALL-1 and APRIL antagonists are used to
 CC block the interaction between APRIL and TALL-1 with TACI or BCMA.
 CC They are useful for treating a mammal suffering from cancer such
 CC as leukaemia, lymphoma, myeloma, cancers of lung and colon and
 CC autoimmune diseases e.g. rheumatoid arthritis, multiple sclerosis,
 CC psoriasis and lupus erythematosus. The present sequence is human
 CC TACI splice variant protein.
 XX
 SQ Sequence 265 AA;

Query Match 100.0%; Score 909; DB 22; Length 265;
 Best Local Similarity 100.0%; Pred. No. 4.1e-83;
 Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MSGLGRSRGRSRVDOERFPQGLWTGVAMRSCPEEQYWPDLTGTCMSCKTICNHQSQR 60
 DB 1 msglgrsrgrsrvdqeerfpqglwtgvamrscpeeqywdplgtcmstctcnhqsr 60
 QY 61 TCAAFCSLSCKRKGKGYDHLRLDCISCSICGQHPKQCAFCENKURSPVNPPELRR 120
 DB 61 tcaafcslsckrkggkfydhllrdciscaslcgghpkqcafcenklsrsvnpnlpeirr 120
 QY 121 QRSGEVENNSDNGRYQGLEHRGSEASPALPGLKLSADQVALVYST 166
 DB 121 qrsgevennsdngryqglehrgseaspalp9lklksadqvalvyst 166

RESULT 3
 AAW75783
 ID AAW75783 standard; Protein; 293 AA.
 XX
 AC AAW75783;
 XX
 DT 18-JAN-1999 (first entry)
 XX
 DE Human lymphocyte surface receptor TACI.
 XX
 KW TACI; transmembrane activator and CAML-interactor;
 KW calcium signal-modulating cyclophilin ligand; human;
 KW lymphocyte surface receptor; human; B-cell; B lymphocyte;
 KW infection; cancer; rheumatoid arthritis; autoimmune disease;
 KW glomerulonephritis; immunosuppressive; graft versus host disease;
 KW transplant rejection; therapy.
 XX
 OS Homo sapiens.
 XX
 PH Location/Qualifiers
 FT Domain 1..166
 FT /label= Extracellular_domain
 FT /note= "Claim 8"
 FT Domain 167..186
 FT /label= Transmembrane_domain
 FT Domain 187..294


```

FT FT /label= Cytoplasmic_domain
FT FT /note= "Claim 6"
FT FT 34..71
FT FT /note= "TNFR_NGFR motif"
XX XX
PN W09839361-A1..
XX XX
PD 11-SEP-1998.
XX XX
XX 03-MAR-1998; 98WO-US04270.
XX XX
XX 03-MAR-1997; 97US-0810572.
XX XX
PA (SUUD-) ST JUDE CHILDREN'S RES HOSPITAL.
XX XX
XX Bram RJ, Von Bulow G;
XX XX
XX WPI; 1998-506346/43.
DR N-PSDB; AAV57328.
XX XX
XX New isolated transmembrane activator protein - used to develop
PT products for treating e.g. infections, cancers, autoimmune and
PT inflammatory conditions, transplant rejection or graft-versus-host
PT disease
XX XX
XX Claim 20; Fig 2a; 89pp; English.
XX XX
XX This is the amino acid sequence of novel human transmembrane
CC activator and CAML-interactor (TACI) protein, a lymphocyte receptor
CC protein that is involved in the calcium activation pathway. TACI
CC is normally present in B-lymphocytes, and to a much lesser extent
CC in immature T-lymphocytes, and can therefore be targetted to
CC specifically regulate B cell responses without affecting T cell
CC activity. TACI cDNA (seeV57328) was isolated from a B-lymphocyte
CC cDNA library using a yeast two-hybrid assay. Also claimed are
CC the C-terminal (see AA75784) and N-terminal (see AA75785) fragments
CC of TACI, recombinant DNA constructs, unicellular hosts, and
CC antibodies to TACI protein. Methods are claimed for identifying a
CC ligand for TACI and for identifying immunosuppressive drugs that
CC selectively block the action of B lymphocytes without affecting
CC mature T lymphocytes. TACI can be activated to increase immune
CC system activity, e.g. for treating infections or cancers. It can
CC be blocked to provide immunosuppression, e.g. for treating
CC autoimmune and inflammatory conditions such as immune complex-
CC induced vasculitis, glomerulonephritis, haemolytic anaemia,
CC myasthenia gravis, type II collagen-induced arthritis, experimental
CC allergic and hyperacute xenograft rejection, rheumatoid arthritis,
CC systemic lupus erythematosus, transplant rejection, cancer or
CC graft versus host disease.
XX XX
XX Sequence 293 AA;
XX SQ
Query Match 100.0%; Score 909; DB 19; Length 293;
Best Local Similarity 100.0%; Pred. No. 4.6e-83;
Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSGLGRRRGRRSRVDQERFPQGLWTGVAMRSCPEEQYWDPLLGTCMSCKTICNHQSQR 60
Db 1 msglgrsrrgrsrvdqeerpqglwtgvamrscpeeeywdpllgtcmsckticnhqsqr 60
QY 61 TCFAFCSLSCKRKQGFYDHLRLDCISCSASICGQHPKQCAFCENKLRSPVNLPPELRR 120
Db 61 tcaafcrslscrkeqgfydhllrdciscasicgqhpqkqayfcenklrspvnlppeirr 120
QY 121 QRSGEVENNSNGRYGQLEHRGSEASPALPGLKLSADQVALVYST 166
Db 121 qrsgevennsnsgrygglehrgeaspaalpglklisadqvalvyst 166
RESULT 4
AAB36312
ID AAB36312 standard; Protein; 293 AA.

```

```

XX AAB36312;
XX AC
XX DT
XX 26-FEB-2001 (first entry)
XX DE
XX Human neutrokinine-alpha binding protein TR17 SEQ ID NO:2.
XX KW
XX Human; neutrokinine-alpha binding protein; NAR protein; TR17; cytostatic;
XX immunosuppressive; neutrotropic; neuroprotective; antiviral; antiallergic;
XX hepatotropic; antidiabetic; antiinflammatory; antitumor; cardiant;
XX ophthalmological; gene therapy; immunodeficiency disorder; diagnosis;
XX autoimmune disorder.
XX KW
XX Homo sapiens.
XX OS
XX WO200058362-A1.
XX PN
XX 05-OCT-2000.
XX PD
XX 24-MAR-2000; 2000WO-US07966.
XX PF
XX 26-MAR-1999; 99US-0126599.
XX PR
XX 10-MAR-2000; 2000US-0188208.
XX PR
XX (HUMA-) HUMAN GENOME SCI INC.
XX PA
XX Ruben SM, Ullrich S, Baker K;
XX PI
XX WPI; 2000-602359/57.
XX DR
XX N-PSDB; AAC84602.
XX XX
XX Nucleic acid encoding a neutrokinine-alpha receptor (NAR) such as TR17,
XX useful for producing TR17 protein which is used in the treatment and
XX diagnosis of autoimmune and immunodeficiency disorders -
XX Claim 1; Fig 1; 398pp; English.
XX PS
XX The present sequence represents the human neutrokinine-alpha binding (NAR)
XX protein designated TR17. TR17 has cytostatic, immunosuppressive,
XX neutrotropic, neuroprotective, antiviral, antiallergic, hepatotropic,
XX antidiabetic, antiinflammatory, antitumor, cardiant and ophthalmological
XX activities and can be used in gene therapy. The TR17 protein and
XX antibodies are useful for treating and diagnosing immunodeficiency
XX disorders and autoimmune disorders. The TR17 polypeptides,
XX polynucleotides, antibodies, agonists and/or antagonists are used for
XX treating various other diseases defined in the specification and as
XX research tools for studying the phenotypic effects that result from
XX inhibiting TR17/TR17 ligand interactions on various cell types.
XX SQ
Sequence 293 AA;
Query Match 100.0%; Score 909; DB 21; Length 293;
Best Local Similarity 100.0%; Pred. No. 4.6e-83;
Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSGLGRRRGRRSRVDQERFPQGLWTGVAMRSCPEEQYWDPLLGTCMSCKTICNHQSQR 60
Db 1 msglgrsrrgrsrvdqeerpqglwtgvamrscpeeeywdpllgtcmsckticnhqsqr 60
QY 61 TCFAFCSLSCKRKQGFYDHLRLDCISCSASICGQHPKQCAFCENKLRSPVNLPPELRR 120
Db 61 tcaafcrslscrkeqgfydhllrdciscasicgqhpqkqayfcenklrspvnlppeirr 120
QY 121 QRSGEVENNSNGRYGQLEHRGSEASPALPGLKLSADQVALVYST 166
Db 121 qrsgevennsnsgrygglehrgeaspaalpglklisadqvalvyst 166
RESULT 5
AAY94000
ID AAY94000 standard; Protein; 293 AA.
XX XX

```

AC AAY94000;
 XX 20-OCT-2000 (first entry)
 DT
 DE A transmembrane activator and CAML-interactor (TACI).
 XX
 KW Human; BR43x2; TACI receptor; extracellular domain; BCMA; B cell protein;
 KW transmembrane activator and CAML-interactor; tumour necrosis factor; TNF;
 KW ztnf4 activity; antibody production; autoimmune disease; amyloidosis;
 KW systemic lupus erythematosus; myasthenia gravis; multiple sclerosis;
 KW rheumatoid arthritis; asthma; bronchitis; emphysema; pyelonephritis;
 KW end stage renal failure; glomerulonephritis; vasculitis; nephritis;
 KW renal neoplasms; multiple myeloma; lymphoma; light chain neuropathy;
 KW immune response; immunosuppression; graft rejection; joint pain;
 KW graft versus host disease; inflammation; swelling; anaemia; septic shock;
 KW insulin dependent diabetes mellitus; Crohn's disease; hypertension;
 KW renal artery stenosis; occlusion; cholesterol; renal emboli.
 XX
 OS Homo sapiens.
 XX
 PN WO200040716-A2.
 XX
 PD 13-JUL-2000.
 XX
 PF 07-JAN-2000; 2000WO-US00396.
 XX
 PR 07-JAN-1999; 99US-0226533.
 XX
 PA (ZYMO) ZYMOGENETICS INC.
 XX
 PI Gross JA, Xu W, Madden K, Yee DP;
 XX WPI; 2000-452538/39.
 DR N-PSDB; AAA58558.
 XX
 PT Inhibiting ztnf4 activity in a mammal, to treat autoimmune diseases,
 PT renal disease, graft versus host disease, and inflammation, comprises
 PT administering a BR43x2, TACI or BCMA extracellular domain polypeptide -
 XX
 PS Disclosure; Page 149-150; 175pp; English.
 XX
 CC The present sequence represents a human transmembrane activator and
 CC CAML-interactor (TACI) receptor. TACI is a tumour necrosis factor (TNF)
 CC receptor. The extracellular domains of BR43x2 (an isoform of TACI), TACI
 CC or BCMA (a related B cell protein) receptor contain a cysteine rich
 CC domain, and are used for inhibiting ztnf4 activity. ztnf4 is a TNF
 CC ligand. They may also be used for inhibiting BR43x2, TACI or BCMA
 CC receptor-ligand engagement associated with activated or resting B
 CC lymphocytes, effector T-cells, or with antibody production. The
 CC antibody production is associated with an autoimmune disease selected
 CC from systemic lupus erythematosus, myasthenia gravis, multiple sclerosis
 CC and rheumatoid arthritis. The ztnf4 activity and BR43x2, TACI or BCMA
 CC receptor-ligand engagement is associated with asthma, bronchitis,
 CC emphysema, end stage renal failure, glomerulonephritis, vasculitis,
 CC nephritis, pyelonephritis, renal neoplasms, multiple myelomas, lymphomas,
 CC light chain neuropathy amyloidosis, moderating immune response,
 CC immunosuppression, graft rejection, graft versus host disease,
 CC inflammation, insulin dependent diabetes mellitus, Crohn's disease, joint
 CC pain, swelling, anaemia, or septic shock. BR43x2, TACI, and BCMA
 CC polypeptides, fusions, antibodies, agonists or antagonists can be used
 CC to treat hypertension, renal artery stenosis, or occlusion, and
 CC cholesterol or renal emboli.
 XX
 SQ Sequence 293 AA;

Query Match 100.0%; Score 909; DB 21; Length 293;
 Best Local Similarity 100.0%; Pred. No. 4.6e-83;
 Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY *1 MSGLGRSRGRSRVDQERFPQGLWTGVAMRSCPEEQYWDPLLTGTCMSCKTICNHQSQR 60
 DB 1 msglgrsrgrsrvdqerfpqglwtgvamrscpeeqywdplltgctmsckticnhqsqr 60

QY 61 TCAAFCSRSLSCRKEQKFDYHLLRDCISCASTCGQHPKQCAFECNKLRSPVNLPELR 120
 DB 61 tcaafcsrslscrkegkfydhllrdciscastcgqhpqkqcfecnklsrpsvnlpeirr 120
 QY 121 QRSGEVNNSDNSGRYQGLEHRGSEASPALPGLKLSADQVALVYST 166
 DB 121 qrsgevennsdngsryqgglehrgeaspalpglklisadqvalvyst 166
 RESULT 6
 AAE09240
 ID AAE09240 standard; Protein; 293 AA.
 XX
 AC AAE09240;
 XX
 DT 19-NOV-2001 (first entry)
 XX
 DE Human TACI protein.
 XX
 KW Human; TNF; tumour necrosis factor; TALL-1; APRIL; TNF receptor;
 KW TNFR; TACI; BCMA; therapy; cancer; leukaemia; myeloma; lymphoma;
 KW autoimmune disease; rheumatoid arthritis; multiple sclerosis;
 KW psoriasis.
 XX
 OS Homo sapiens.
 XX
 PN WO200160397-A1.
 XX
 PD 23-AUG-2001.
 XX
 PF 28-NOV-2000; 2000WO-US2378.
 XX
 PR 16-FEB-2000; 2000US-0182938.
 PR 22-AUG-2000; 2000US-0226986.
 XX
 PA (GETH) GENENTECH INC.
 XX
 PI Ashkenazi AJ, Dodge KH, Grewal I, Kim KJ, Marsters SA, Pitti RM;
 PI Yan M;
 XX WPI; 2001-541628/60.
 DR N-PSDB; AAD15901.
 XX
 PT Inhibiting or neutralizing TALL-1 or APRIL polypeptide biological
 PT activity, for treating autoimmune disorders and cancer, comprises
 PT exposing the cells to TALL-1 or APRIL polypeptide agonists or
 PT antagonists -
 XX
 PS Example 1; Fig 1; 160pp; English.
 XX
 CC The invention relates to methods of using one or more agonists or
 CC antagonists to modulate the activity of the members of TNF (tumour
 CC necrosis factor) especially TALL-1, APRIL and TNF receptor (TNFR)
 CC e.g. TACI or BCMA. The method is useful for treating pathological
 CC conditions or diseases associated with increased TALL-1 and APRIL
 CC expression or activity. TALL-1 and APRIL antagonists are used to
 CC block the interaction between APRIL and TALL-1 with TACI or BCMA.
 CC They are useful for treating a mammal suffering from cancer such
 CC as leukaemia, lymphoma, myeloma, cancers of lung and colon and
 CC autoimmune diseases e.g. rheumatoid arthritis, multiple sclerosis,
 CC psoriasis and lupus erythematosus. The present sequence is human
 CC TACI protein.
 XX
 SQ Sequence 293 AA;

Query Match 100.0%; Score 909; DB 22; Length 293;
 Best Local Similarity 100.0%; Pred. No. 4.6e-83;
 Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MSGLGRSRGRSRVDQERFPQGLWTGVAMRSCPEEQYWDPLLTGTCMSCKTICNHQSQR 60
 DB 1 msglgrsrgrsrvdqerfpqglwtgvamrscpeeqywdplltgctmsckticnhqsqr 60

Db 1 msglgrsrrgrsvdqeefpqlwtgvmrscpeeqywdpllgctmsckticnhsqr 60
 Qy 61 TCAAFCSLSCKRKGKGFYDHLRDCISCSICGQHPKQCAFCECNKLRSPVNLPPELRR 120
 Db 61 tcaafcsrlscrkeqgkfydhllrdciscaslcgqhpqkqcfayfcenklrsvnlppelrr 120
 Qy 121 QRSGEVNNSDNSGRYQGLEHRGSEASPALPGLKLSADQVALVYST 166
 Db 121 qrsgevennsdngryqglehrgeaspalpglksadqvalvyst 166
 RESULT 7
 AAY71914
 ID AAY71914 standard; Protein: 293 AA.
 XX
 AC AAY71914;
 XX
 DT 26-MAR-2001 (first entry)
 XX
 DE Human tumour necrosis factor receptor (TACI) protein.
 XX
 KW Human; transmembrane activator and CAML interactor; TACI;
 KW tumour necrosis factor receptor; TNF; autoimmune disease; diabetes;
 KW calcium-signal modulating cyclophilin ligand; CAML; viral infection;
 KW neutroline alpha polypeptide; TACI-Ligand; TACI-L; cytostatic; therapy;
 KW neuroprotective; antidiabetic; antiviral; antiinflammatory; tumour;
 KW antiarthritic; antirheumatic; immunosuppressive; multiple sclerosis;
 KW rheumatoid arthritis; graft rejection; inflammation; cell proliferation;
 KW cell death; immunoglobulin E-mediated allergic reaction; IgE.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 DE Domain 2..166
 FT /label= Extracellular domain
 FT /note= "Binds with amino acids 123-285 of extracellular
 FT domain of TACI-L"
 XX
 PN W0200067034-A1.
 XX
 XX 09-NOV-2000.
 XX
 XX 14-APR-2000; 2000WO-US10282.
 XX
 XX 30-APR-1999; 99US-0302863.
 XX
 XX (IMMU) IMMUNEX CORP.
 XX
 PI Goodwin RG, Din WS;
 XX
 XX WPI; 2001-016005/02.
 XX
 XX N-PSDB; AAD02006.
 XX
 XX Use of new interactions between tumour necrosis factor receptors (TACI)
 XX and TACI ligands to screen candidate molecules for determining agonist
 XX and antagonist interactions which are used for treating inflammation -
 XX
 XX Claim 10; Fig 1b; 46pp; English.
 XX
 XX The present sequence is a human tumour necrosis factor receptor (TACI)
 XX protein. TACI (transmembrane activator and calcium-signal modulating
 XX cyclophilin ligand (CAML)-interactor) forms a complex with neutroline
 XX alpha polypeptide (TACI-Ligand). The antagonist or agonist of
 XX TACI/TACI-L complex is useful for modulating an intracellular signalling
 XX cascade mediated by TACI/TACI-L complex. Antagonists of TACI/TACI-L
 XX complex are used to inhibit the interaction between TACI and TACI-L for
 XX therapeutic purposes to treat tumour and tumour metastasis and to combat
 XX various autoimmune diseases e.g. multiple sclerosis and diabetes, as
 XX well as other disorders, such as viral infection, rheumatoid arthritis,
 XX graft rejection, and immunoglobulin (Ig) E-mediated allergic reactions
 XX and inflammation. The interaction is used to study cellular processes
 XX associated with tumour necrosis factor (TNF)-receptors such as immune
 XX regulation, cell proliferation, cell death and inflammatory responses.

CC The interaction between the extracellular region of TACI and TACI-L can
 CC be used to further develop understanding of which cell types TACI-L
 CC acts upon.
 XX
 SQ Sequence 293 AA;
 Query Match 100.0%; Score 909; DB 22; Length 293;
 Best Local Similarity 100.0%; Pred. No. 4.6e-83;
 Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 MSGLGRSRRGRSVRVDQERFPQGLWTGVMRSCPEEQYWDPLLGCTMSCKTICNHQSQR 60
 Db 1 msglgrsrrgrsvrdqeefpqlwtgvmrscpeeqywdpllgctmsckticnhsqr 60
 Qy 61 TCAAFCSLSCKRKGKGFYDHLRDCISCSICGQHPKQCAFCECNKLRSPVNLPPELRR 120
 Db 61 tcaafcsrlscrkeqgkfydhllrdciscaslcgqhpqkqcfayfcenklrsvnlppelrr 120
 Qy 121 QRSGEVNNSDNSGRYQGLEHRGSEASPALPGLKLSADQVALVYST 166
 Db 121 qrsgevennsdngryqglehrgeaspalpglksadqvalvyst 166
 RESULT 8
 AAY93998
 ID AAY93998 standard; Protein: 247 AA.
 XX
 AC AAY93998;
 XX
 DT 20-OCT-2000 (first entry)
 XX
 DE Human BR43x2, an isoform of the TACI receptor.
 XX
 KW Human; BR43x2; TACI receptor; extracellular domain; BCMA; B cell protein;
 KW transmembrane activator and CAML-interactor; tumour necrosis factor; TNF;
 KW ztnf4 activity; antibody production; autoimmune disease; amyloidosis;
 KW systemic lupus erythematosus; myasthenia gravis; multiple sclerosis;
 KW rheumatoid arthritis; asthma; bronchitis; emphysema; pyelonephritis;
 KW end stage renal failure; glomerulonephritis; vasculitis; nephritis;
 KW renal neoplasm; multiple myeloma; lymphoma; light chain neuropathy;
 KW immune response; immunosuppression; graft rejection; joint pain;
 KW graft versus host disease; inflammation; swelling; anaemia; septic shock;
 KW insulin dependent diabetes mellitus; Crohn's disease; hypertension;
 KW renal artery stenosis; occlusion; cholesterol; renal emboli.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 DE Domain 1..120
 FT /note= "extracellular domain"
 FT Region 25..58
 FT /note= "cysteine-rich pseudo repeat"
 FT Domain 121..133
 FT /note= "transmembrane domain"
 FT Domain 134..247
 FT /note= "cytoplasmic domain"
 XX
 PN W0200040716-A2.
 XX
 XX 13-JUL-2000.
 XX
 XX 07-JAN-2000; 2000WO-US00396.
 XX
 XX 07-JAN-1999; 99US-0226533.
 XX
 XX (ZYMO) ZYMOGENETICS INC.
 XX
 XX Gross JA, Xu W, Madden K, Yee DP;
 XX WPI; 2000-452538/39.
 XX
 XX N-PSDB; AAA58556.
 XX

KW inflammatory bowel disease; septic shock; ulcerative colitis;
KW Crohn's disease; chronic myelogenous leukemia; cancer; liver
KW disease; Hodgkin's disease; osteoarthritis; Lyme's disease;
KW cachexia; autoimmune disease; myasthenia gravis; autoimmune diabetes;
KW systemic lupus erythematosus; transgenic animal; diagnosis;
KW prognosis; prophylactic; therapeutic; human.
XX Homo sapiens.
XX WO200039284-A1.
XX 06-JUL-2000.
XX 23-DEC-1999; 99WO-US31025.
XX 30-DEC-1998; 98US-0223546.
XX (MILL-) MILLENNIUM PHARM INC.
XX Holtzman DA;
XX WPI: 2000-465743/40.
XX N-PSDB; AAA47454.
XX Novel nucleic acid sequences encoding TANGO-128, 140, 197, 212, 213,
PT 224 and 239 polypeptides useful for the treatment of asthma, rheumatoid
PT arthritis, psoriasis and autoimmune diseases
XX Claim 8; Fig 3; 209pp; English.
XX Nucleic acids encoding TANGO polypeptides are useful as modulating
CC agents for regulating cellular processes like asthma, graft
CC versus-host diseases, rheumatoid arthritis, psoriasis, inflammatory
CC bowel disease, septic shock, ulcerative colitis, Crohn's disease,
CC chronic myelogenous leukemia, cancer, liver disease, Hodgkin's
CC disease, osteoarthritis, Lyme's disease, cachexia and autoimmune
CC diseases e.g. myasthenia gravis, autoimmune diabetes and systemic
CC lupus erythematosus. The nucleic acids are also useful for producing
CC transgenic animals and the TANGO polypeptides themselves. Partial
CC TANGO-128, 140, 197, 212, 213, 224, 239 sequences are useful in
CC forensic biology, for diagnostic assays, prognostic assays,
CC pharmacogenomics and for monitoring clinical trials. TANGO
CC polypeptides are suitable for both prophylactic and therapeutic
CC methods for treating a subject at risk of a disorder or having a
CC disorder associated with aberrant TANGO expression. A wide range
CC of cellular disorders can be treated.
XX Sequence 197 AA;

Query Match 10.5%; Score 95; DB 21; Length 197;
Best Local Similarity 22.4%; Pred. No. 0.13;
Matches 28; Conservative 20; Mismatches 35; Indels 42; Gaps 7;
QY 27 TGVAMRSCPEQYWDPLLTGTCMCKTIC--NHQSORTCA-----AFCRSLSCRKEGKF 78
Db 18 tglptmdcgeneywdq-wgrcvctqr-cgpgqelskdcgygedgdayctacpprrykw 75
QY 79 YDHLRLDCISCA-----SICGQHPKQCAVFCENKLRSPVNLPPELRRORS 123
Db 76 ghkqcqcitcavinrvqkvinctatsnavcqd-----c-----lprfyrktri 118
QY 124 GEVEN 128
Db 119 gglqd 123

RESULT 11
ID AAA42025
XX AAA42025 standard; Protein; 1589 AA.
AC AAA42025;
XX

DT 22-OCT-2001 (first entry)
XX Human polypeptide SEQ ID NO 6956.
XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW Leukaemia.
XX Homo sapiens.
XX WO200153312-A1.
XX 26-JUL-2001.
XX 26-DEC-2000; 2000WO-US34263.
XX 21-JAN-2000; 2000US-0488725.
XX 25-APR-2000; 2000US-0552317.
XX 09-JUL-2000; 2000US-0598042.
XX 19-JUL-2000; 2000US-0620312.
XX 03-AUG-2000; 2000US-0653450.
XX 14-SEP-2000; 2000US-0662191.
XX 19-OCT-2000; 2000US-0693036.
XX 29-NOV-2000; 2000US-0727344.
XX (HYSE-) HYSEQ INC.
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX WPI: 2001-442253/47.
XX N-PSDB; AAI61181.
XX Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX Example 2; SEQ ID NO 6956; 10078pp; English.
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX Sequence 1589 AA;

Query Match 10.2%; Score 93; DB 22; Length 1589;
Best Local Similarity 25.8%; Pred. No. 2.5;
Matches 40; Conservative 17; Mismatches 68; Indels 30; Gaps 7;
QY 22 PQGLWTCVAMRSCPEQYWDPLLTGTCMCKTIC--CNHQSQRTCA-----AFCRSLSCR 72
Db 1268 pegtteitsyeitprirvwrqtlrcrsaaqvciclgqlersiaeksvnkvtc--lvc 1325
QY 73 KEQKGFYDHLRLDCISCAICG---QHPKQCA-----YFC-----ENKLRSPVNLPP 116
Db 1326 kgdn---deflllcdgdrghyhrpkmavpegdwfctvctlaqvegtkqpfk 1382

DR WPI; 1998-583603/49.
DR N-PSDB; AAV68396.
XX
PT Transcriptional regulator gene family containing bromodomain - may
PT be expressed in testis tissue and is useful for treatment of cancer
PT and other proliferative disorders
XX
PS Claim 1; Page 72-88; 187pp; Japanese.
XX
CC This sequence represents the human BAZ2-alpha protein, a member of a
CC family of transcriptional regulator genes containing a bromodomain (BAZ,
CC Bromodomain with Atypical Zinc Finger) which are expressed specifically
CC in testis tissue and also in certain tumour lines. Transgenic cells may
CC be used for the preparation of the BAZ1-alpha, BAZ1-beta, BAZ2-alpha and
CC BAZ2-beta proteins. These proteins can be used in the treatment of cancer
CC and other proliferative disorders, and in screening of compounds for
CC their binding ability to the expression products (e.g. for use as drugs
CC by modulation of transcriptional regulation).
XX
SQ Sequence 1878 AA;

Query Match 9.8%; Score 89; DB 19; Length 1878;
Best Local Similarity 23.3%; Pred. No. 7.5;
Matches 37; Conservative 18; Mismatches 66; Indels 38; Gaps 6;

QY 22 PQGLWTGVAMRSCPEEQYWDPLLGTCTMCKTIC-----NHQSQTCAAFGRS 68
Db 1599 pegttteisyeltprirwrtlqrcrsaaahvclclghlersiaweksvnkvtc----- 1652
QY 69 LSCRKEQKGYDHLRDCISASICG---QHPKQCA-----YFC-----ENKLRSPV 112
Db 1653 lvcrrkgdn---defllldcgdrghychrpkmeavpegdwfctvciaqvgveftgkp 1709
QY 113 NLPPELRQRSGEVENNSDNGRYOGLEHRGSEASPALP 151
Db 1710 gfpkrqgkrksygslnfsegdgrrrrrvllkgrespaagp 1748

Search completed: June 25, 2002, 16:16:09
Job time: 289 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 25, 2002, 16:22:44 ; Search time 27.55 Seconds
(without alignments)
94.164 Million cell updates/sec

Title: US-09-854-864-16

Perfect score: 405

Sequence: 1 CPPEQYNDPLGLTGMCKTI.....DCISCASICGQHPKQCAFC 67

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	76.5	18.9	354	1	VEGD_HUMAN
2	75	18.5	1877	1	PKCS_MOUSE
3	74.5	18.4	969	1	PAC1_HUMAN
4	71.5	17.7	773	1	ITB1_BOVIN
5	71.5	17.7	1680	1	FUR2_DROME
6	70	17.3	1124	1	TIE2_HUMAN
7	69.5	17.2	798	1	ITB1_MOUSE
8	69	17.0	913	1	PKCS_HUMAN
9	68.5	16.9	501	1	TRA2_MOUSE
10	68	16.8	1339	1	ERB3_RAT
11	67.5	16.7	3084	1	LMAL_MOUSE
12	67	16.5	1210	1	EGFR_MOUSE
13	67	16.5	3110	1	LMAL_HUMAN
14	66.5	16.4	184	1	TR17_HUMAN
15	66.5	16.4	798	1	ITB1_FELCA
16	66.5	16.4	937	1	PACA_RAT
17	66	16.3	60	1	MT_ICTPU
18	66	16.3	1122	1	TIE2_MOUSE
19	66	16.3	1210	1	EGFR_HUMAN
20	65.5	16.2	60	1	MT_GADMO
21	65.5	16.2	358	1	VEGD_MOUSE
22	65.5	16.2	1426	1	EGFR_DROME
23	65	16.0	799	1	ITB1_RAT
24	65	16.0	1342	1	ERB3_HUMAN
25	65	16.0	1696	1	PKCS_BRACL
26	65	16.0	1877	1	PKCS_RAT
27	65	16.0	2569	1	LMAL_MOUSE
28	65	16.0	3106	1	LMAL_MOUSE
29	64.5	15.9	443	1	FBL4_MOUSE
30	64.5	15.9	798	1	ITB1_XENLA
31	64	15.8	97	1	ECLH_DROME
32	64	15.8	119	1	ANTA_HAEGH
33	64	15.8	1125	1	TIE2_BOVIN

RESULT	1	VEGD_HUMAN	STANDARD;	PRT;	354 AA.
AC	043915;				
DT	01-MAR-2002 (Rel. 41, Created)				
DT	01-MAR-2002 (Rel. 41, Last sequence update)				
DT	01-MAR-2002 (Rel. 41, Last annotation update)				
DE	Vascular endothelial growth factor D precursor (VEGF-D) (c-fos induced growth factor) (FIGF).				
GN	FIGF OR VEGFD.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Lung;				
RX	MEDLINE=97349118; PubMed=9205122;				
RA	Yamada Y., Nezu J.-I., Shimane M., Hirata Y.;				
RT	"Molecular cloning of a novel vascular endothelial growth factor,				
RT	VEGF-D.";				
RL	Genomics 42:483-488(1997).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Lung;				
RX	MEDLINE=98140120; PubMed=9479493;				
RA	Rocchigiani M., Lestingi M., Luddi A., Orlandini M., Franco B.,				
RA	Rossi E., Ballabio A., Zuffardi O., Oliviero S.;				
RT	"Human FIGF: cloning, gene structure, and mapping to chromosome Xp22.1				
RT	between the FIGA and the GRPR genes.";				
RL	Genomics 47:207-216(1998).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=98118549; PubMed=9435229;				
RA	Achen M.G., Jeltsch M., Kuk E., Maekinen T., Vitali A., Wilks A.F.,				
RA	Alitalo K., Stacker S.A.;				
RT	"Vascular endothelial growth factor D (VEGF-D) is a ligand for the				
RT	tyrosine kinases VEGF receptor 2 (Flk1) and VEGF receptor 3 (Flt4).";				
RL	Proc. Natl. Acad. Sci. U.S.A. 95:548-553(1998).				
RN	[4]				
RP	PROCESSING, AND SEQUENCE OF 89-94; 100-105 AND 206-213.				
RX	MEDLINE=20011413; PubMed=10542248;				
RA	Stacker S.A., Stenvers K.L., Caesar C., Vitali A., Domagala T.,				
RA	Nice E.C., Roufail S., Simpson R.J., Moritz R., Karpanen T.,				
RT	Alitalo K., Achen M.G.;				
RT	"Biosynthesis of vascular endothelial growth factor-D involves				
RL	proteolytic processing which generates non-covalent homodimers.";				
RL	J. Biol. Chem. 274:32127-32136(1999).				
CC	-!- FUNCTION: Growth factor active in angiogenesis lymphangiogenesis				
CC	and endothelial cell growth, stimulating their proliferation and				
CC	migration and also has effects on the permeability of blood				
CC	vessels. May function in the formation of the venous and lymphatic				
CC	vascular systems during embryogenesis, and also in the maintenance				
CC	of differentiated lymphatic endothelium in adults. Binds and				
CC	activates VEGFR-2 (Flk1) and VEGFR-3 (Flt4) receptors.				
CC	-!- SUBUNIT: Homodimer; non-covalent and antiparallel.				

Q13639 homo sapien
Q25092 hirudo medi
Q28833 sus scrofa
P25119 mus musculus
P71397 haemophilus
P52721 thermarces
O70528 cavia porce
Q93038 h wsl-1 pro
P12607 xenopus lae
P05556 homo sapien
P07228 gallus gall
Q9ubw7 homo sapien

ALIGNMENTS

RC TISSUE-Placenta;
RX MEDLINE-94235049; PubMed-8179631;
RA Tsuji A., Higashine K., Hine C., Mori K., Tamai Y., Nagamune H.,
RA Matsuda Y.;
RT "Identification of novel cDNAs encoding human kexin-like protease,
RT PACE4 isoforms.";
RL Biochem. Biophys. Res. Commun. 200:943-950(1994).
RN [3]
RP ERRATUM.
RX MEDLINE-95071480; PubMed-7980617;
RA Tsuji A., Higashine K., Hine C., Mori K., Tamai Y., Nagamune H.,
RA Matsuda Y.;
RT "Identification of novel cDNAs encoding human kexin-like protease,
RT PACE4 isoforms.";
RL Biochem. Biophys. Res. Commun. 204:1381-1382(1994).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM PACE4A-II).
RC TISSUE-Placenta;
RA Mori K., Imamaki A., Kii S., Nagamune H., Nagahama M., Tsuji A.,
RA Matsuda Y.;
RT "Identification of a novel PACE4 isoform, PACE4E.";
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A. (ISOFORMS PACE4E-I AND PACE4E-II).
RC TISSUE-Cerebellum;
RX MEDLINE-97335942; PubMed-9192737;
RA Mori K., Kii S., Tsuji A., Nagahama M., Imamaki A., Hayashi K.,
RA Akamatsu T., Nagamune H., Matsuda Y.;
RT "A novel human PACE4 isoform, PACE4E is an active processing protease
RT containing a hydrophobic cluster at the carboxy terminus.";
RL J. Biochem. 121:941-948(1997).
RN [6]
RP SEQUENCE FROM N.A. (ISOFORMS PACE4A-I; A-II; CS; D; E-I; E-II).
RX MEDLINE-98021085; PubMed-9378725;
RA Tsuji A., Hine C., Tamai Y., Yonemoto K., Mori K., Yoshida S.,
RA Bando M., Sakai E., Mori K., Akamatsu T., Matsuda Y.;
RT "Genomic organization and alternative splicing of human PACE4 (SPC4),
RT kexin-like processing endoprotease.";
RL J. Biochem. 122:438-452(1997).
RN [7]
RP ALTERNATIVE SPLICING (ISOFORM PACE4CS).
RX MEDLINE-97064242; PubMed-8906861;
RA Zhong M., Benjannet S., Lazure C., Munzer S., Seidah N.G.;
RT "Functional analysis of human PACE4-A and PACE4-C isoforms:
RT identification of a new PACE4-CS isoform.";
RL FEBS Lett. 396:31-36(1996).
RN [8]
RP CHARACTERIZATION.
RX MEDLINE-99233559; PubMed-10215603;
RA Sucic J.F., Moehring J.M., Incencio N.M., Luchini J.W.,
RA Moehring T.J.;
RT "Endoprotease PACE4 is Ca2+-dependent and temperature-sensitive and
RT can partly rescue the phenotype of a furin-deficient cell strain.";
RL Biochem. J. 339:639-647(1999).
RN [9]
RP PROCESSING.
RX MEDLINE-98408849; PubMed-9738469;
RA Nagahama M., Taniguchi T., Hashimoto E., Imamaki A., Mori K.,
RA Tsuji A., Matsuda Y.;
RT "Biochemical processing and quaternary interactions of proprotein
RT convertase SPC4 (PACE4).";
RL FEBS Lett. 434:155-159(1998).
CC -1- FUNCTION: LIKELY TO REPRESENT AN ENDOPROTEASE ACTIVITY WITHIN THE
CC CONSTITUTIVE SECRETORY PATHWAY, WITH UNIQUE RESTRICTED
CC DISTRIBUTION IN BOTH NEUROENDOCRINE AND NON-NEUROENDOCRINE TISSUES
CC AND CAPABLE OF CLEAVAGE AT THE RX(K/R)R CONSENSUS MOTIF.
CC -1- CATALYTIC ACTIVITY: RELEASE OF NATURE PROTEINS FROM THEIR
CC PROPROTEINS BY CLEAVAGE OF ARG-XAA-YAA-ARG-[ZAA BONDS,
CC WHERE XAA CAN BE ANY AMINO ACID AND YAA IS ARG OR LYS.
CC -1- COFACTOR: PACE4A IS PROBABLY CALCIUM-DEPENDENT.
CC -1- SUBUNIT: THE PACE4A-I PRECURSOR PROTEIN SEEMS TO EXIST IN THE
CC RETICULUM ENDOPLASMIC AS BOTH A MONOMER AND A DIMER-SIZED COMPLEX
CC WHEREAS NATURE PACE4A-I EXISTS ONLY AS A MONOMER, SUGGESTING THAT

CC CC PROPEPTIDE CLEAVAGE AFFECTS ITS TERTIARY OR QUATERNARY STRUCTURE.
CC CC -1- SUBCELLULAR LOCATION: PACE4A-I AND PACE4A-II ARE SECRETED. PACE4C
CC CC AND PACE4S ARE NOT SECRETED AND REMAIN PROBABLY IN ZYMOGEN FORM
CC CC IN ENDOPLASMIC RETICULUM. PACE4E-I AND PACE4E-II ARE RETAINED
CC CC INTRACELLULARLY PROBABLY THROUGH A HYDROPHOBIC CLUSTER IN THEIR C-
CC CC TERMINUS. PACE4B MIGHT BE SECRETED.
CC CC -1- ALTERNATIVE PRODUCTS: 8 ISOFORMS: PACE4A-I/PACE4 (SHOWN HERE),
CC CC PACE4A-II, PACE4B/PACE4.1, PACE4C, PACE4CS, PACE4E-I AND
CC CC PACE4E-II; ARE PRODUCED BY ALTERNATIVE SPLICING. ISOFORMS PACE4B,
CC CC C. CS AND D MIGHT BE ENZYMATICALLY INACTIVE.
CC CC -1- TISSUE SPECIFICITY: EACH PACE4 ISOFORM EXHIBITS A UNIQUE
CC CC RESTRICTED DISTRIBUTION. PACE4A-I IS EXPRESSED IN HEART, BRAIN,
CC CC PLACENTA, LUNG, SKELETAL MUSCLE, KIDNEY, PANCREAS, BUT AT
CC CC COMPARATIVELY HIGHER LEVELS IN THE LIVER. PACE4A-II IS AT LEAST
CC CC EXPRESSED IN PLACENTA. PACE4B WAS ONLY FOUND IN THE EMBRYONIC
CC CC KIDNEY CELL LINE FROM WHICH IT WAS ISOLATED. PACE4C AND PACE4D ARE
CC CC EXPRESSED IN PLACENTA. PACE4E-I IS EXPRESSED IN CEREBELLUM,
CC CC PLACENTA AND PITUITARY. PACE4E-II IS AT LEAST PRESENT IN
CC CC CEREBELLUM.
CC CC -1- DOMAIN: THE PROPEPTIDE DOMAIN ACTS AS AN INTRAMOLECULAR CHAPERONE
CC CC ASSISTING THE FOLDING OF THE ZYMOGEN WITHIN THE ENDOPLASMIC
CC CC RETICULUM. ISOFORM PACE4D LACKS THE PROPEPTIDE DOMAIN.
CC CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8; ALSO KNOWN AS THE
CC CC SUBTILASE FAMILY.
CC CC -1- SIMILARITY: CONTAINS 1 HOMO B/P DOMAIN.
CC CC -----
CC CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC CC EMBL: M80482; AAA59998.1;
CC CC EMBL: AB001914; BAA21620.1;
CC CC EMBL: AB001898; BAA21620.1; JOINED.
CC CC EMBL: AB001900; BAA21620.1; JOINED.
CC CC EMBL: AB001901; BAA21620.1; JOINED.
CC CC EMBL: AB001902; BAA21620.1; JOINED.
CC CC EMBL: AB001903; BAA21620.1; JOINED.
CC CC EMBL: AB001904; BAA21620.1; JOINED.
CC CC EMBL: AB001905; BAA21620.1; JOINED.
CC CC EMBL: AB001914; BAA21621.1;
CC CC EMBL: AB001898; BAA21621.1; JOINED.
CC CC EMBL: AB001900; BAA21621.1; JOINED.
CC CC EMBL: AB001901; BAA21621.1; JOINED.
CC CC EMBL: AB001902; BAA21621.1; JOINED.
CC CC EMBL: AB001903; BAA21621.1; JOINED.
CC CC EMBL: AB001904; BAA21621.1; JOINED.
CC CC EMBL: AB001905; BAA21621.1; JOINED.
CC CC EMBL: AB001906; BAA21621.1; JOINED.
CC CC EMBL: AB001907; BAA21621.1; JOINED.
CC CC EMBL: AB001908; BAA21621.1; JOINED.
CC CC EMBL: AB001909; BAA21621.1; JOINED.
CC CC EMBL: AB001914; BAA21622.1;
CC CC EMBL: AB001901; BAA21622.1; JOINED.
CC CC EMBL: AB001902; BAA21622.1; JOINED.
CC CC EMBL: AB001903; BAA21622.1; JOINED.
CC CC EMBL: AB001904; BAA21622.1; JOINED.
CC CC EMBL: AB001905; BAA21622.1; JOINED.
CC CC EMBL: AB001906; BAA21622.1; JOINED.
CC CC EMBL: AB001907; BAA21622.1; JOINED.
CC CC EMBL: AB001908; BAA21622.1; JOINED.
CC CC EMBL: AB001914; BAA21623.1;
CC CC EMBL: AB001898; BAA21623.1; JOINED.
CC CC EMBL: AB001900; BAA21623.1; JOINED.
CC CC EMBL: AB001901; BAA21623.1; JOINED.
CC CC EMBL: AB001902; BAA21623.1; JOINED.
CC CC EMBL: AB001903; BAA21623.1; JOINED.
CC CC EMBL: AB001904; BAA21623.1; JOINED.
CC CC EMBL: AB001905; BAA21623.1; JOINED.
CC CC EMBL: AB001906; BAA21623.1; JOINED.

DR	EMBL; AB001907; BAA21623.1; JOINED.	CC	FOR VCAM1 AND RECOGNIZES THE SEQUENCE Q-I-D-S IN VCAM1. INTEGRIN
DR	EMBL; AB001908; BAA21623.1; JOINED.	CC	ALPHA-9/BETA-1 IS A RECEPTOR FOR VCAM1, CYTOTACTIN AND
DR	EMBL; AB001909; BAA21623.1; JOINED.	CC	OSTEOPOINTIN. IT RECOGNIZES THE SEQUENCE A-E-I-D-G-E-I-L IN
DR	EMBL; AB001914; BAA21624.1; JOINED.	CC	CYTOTACTIN. INTEGRIN ALPHA-3/BETA-1 IS A RECEPTOR FOR EPILIGRIN
DR	EMBL; AB001898; BAA21624.1; JOINED.	CC	AND THROMBOSPONDIN. INTEGRIN ALPHA-V/BETA-1 IS A RECEPTOR FOR
DR	EMBL; AB001900; BAA21624.1; JOINED.	CC	VITRONECTIN. BETA-1 INTEGRINS RECOGNIZE THE SEQUENCE R-G-D IN A
DR	EMBL; AB001901; BAA21624.1; JOINED.	CC	WIDE ARRAY OF LIGANDS.
DR	EMBL; AB001902; BAA21624.1; JOINED.	CC	-I- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. BETA-1
DR	EMBL; AB001903; BAA21624.1; JOINED.	CC	ASSOCIATES WITH EITHER ALPHA-1, ALPHA-2, ALPHA-3, ALPHA-4, ALPHA-
DR	EMBL; AB001904; BAA21624.1; JOINED.	CC	5, ALPHA-6, ALPHA-7, ALPHA-8, ALPHA-9, ALPHA-10, ALPHA-11 OR
DR	EMBL; AB001905; BAA21624.1; JOINED.	CC	ALPHA-V.
DR	EMBL; AB001906; BAA21624.1; JOINED.	CC	-I- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. ISOFORM BETA-1B
DR	EMBL; AB001907; BAA21624.1; JOINED.	CC	DOES NOT LOCALIZE TO FOCAL ADHESIONS.
DR	EMBL; AB001908; BAA21624.1; JOINED.	CC	-I- PTM: THE CYSTEINE RESIDUES ARE INVOLVED IN INTRACHAIN DISULFIDE
DR	EMBL; AB001910; BAA21624.1; JOINED.	CC	BONDS.
DR	EMBL; AB001911; BAA21624.1; JOINED.	CC	-I- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.
DR	EMBL; AB001912; BAA21624.1; JOINED.	CC	-I- SIMILARITY: CONTAINS 1 VFMA-LIKE DOMAIN.
DR	EMBL; AB001913; BAA21624.1; JOINED.	CC	-----
DR	EMBL; AB001914; BAA21625.1; JOINED.	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
DR	EMBL; AB001898; BAA21625.1; JOINED.	CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
DR	EMBL; AB001900; BAA21625.1; JOINED.	CC	the European Bioinformatics Institute. There are no restrictions on its
DR	EMBL; AB001901; BAA21625.1; JOINED.	CC	use by non-profit institutions as long as its content is in no way
DR	EMBL; AB001902; BAA21625.1; JOINED.	CC	modified and this statement is not removed. Usage by and for commercial
DR	EMBL; AB001903; BAA21625.1; JOINED.	CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
DR	EMBL; AB001904; BAA21625.1; JOINED.	CC	or send an email to license@isb-sib.ch).
DR	EMBL; AB001905; BAA21625.1; JOINED.	CC	-----
DR	EMBL; AB001906; BAA21625.1; JOINED.	CC	EMBL; UI0865; AAA80571.1; -
DR	EMBL; AB001907; BAA21625.1; JOINED.	DR	InterPro: IPR000561; EGF-like.
DR	EMBL; AB001908; BAA21625.1; JOINED.	DR	InterPro: IPR002389; Integrin_beta_C.
DR		DR	InterPro: IPR001169; Integrin_beta_C.
DR		DR	InterPro: IPR003659; PSI.
DR		DR	Pfam: PF00362; Integrin_B; 1.
DR		DR	ProDom: PD001811; Integrin_B; 1.
DR		DR	SMART: SM00001; EGF_like; 1.
DR		DR	SMART: SM00187; INB; 1.
DR		DR	SMART: SM00423; PSI; 1.
DR		DR	PROSITE: PS00243; INTEGRIN_BETA; 3.
DR		DR	PROSITE: PS00022; EGF_1; UNKNOWN_2.
KW	Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;		
KW	Repeat; Phosphorylation.		
FT	NON_TER	1	
FT	DOMAIN	<1	705
FT	TRANSMEM	706	728
FT	DOMAIN	729	>773
FT	DOMAIN	117	355
FT	DOMAIN	443	612
FT	REPEAT	443	492
FT	REPEAT	493	536
FT	REPEAT	537	575
FT	REPEAT	576	612
FT	DISULFID	4	441
FT	DISULFID	12	22
FT	DISULFID	15	52
FT	DISULFID	25	41
FT	DISULFID	184	190
FT	DISULFID	238	278
FT	DISULFID	378	392
FT	DISULFID	412	668
FT	DISULFID	439	443
FT	DISULFID	454	466
FT	DISULFID	463	502
FT	DISULFID	468	477
FT	DISULFID	479	493
FT	DISULFID	508	513
FT	DISULFID	510	545
FT	DISULFID	515	530
FT	DISULFID	532	537
FT	DISULFID	551	556
FT	DISULFID	553	584
FT	DISULFID	558	567
FT	DISULFID	569	576
FT	DISULFID	590	595
FT	DISULFID	592	638

DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Angiotensin 1 receptor precursor (EC 2.7.1.112) (Tyrosine-protein
 DE kinase receptor TIE-2) (Tyrosine-protein kinase receptor TEK) (P140
 DE TEK) (Tunica interna endothelial cell kinase).
 GN TEK OR TIE2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=93173509; PubMed=8382358;
 RA Ziegler S.F., Bird T.A., Schneringer J.A., Schooley K.A., Baum P.R.;
 RT "Molecular cloning and characterization of a novel receptor protein
 RT tyrosine kinase from human placenta.";
 RL Oncogene 8:663-670(1993).
 RN [2]
 RP VARIANT VMC1 TRP-849.
 RX MEDLINE=97134665; PubMed=8980225;
 RA Vikkula M., Boon L.M., Carraway K.L. III, Calvert J.T., Diamonti A.J.,
 RA Goumrov B., Pasyk K.A., Marchuk D.A., Warman M.L., Cantley L.C.,
 RA Mulliken J.B., Olse B.R.;
 RT "Vascular dysmorphogenesis caused by an activating mutation in the
 RT receptor tyrosine kinase TIE2.";
 RL Cell 87:1181-1190(1996).
 RN [3]
 RP VARIANTS VMC1 TRP-849 AND SER-897.
 RX MEDLINE=9299243; PubMed=10369874;
 RA Calvert J.T., Kiley T.J., Kontos C.D., Cha E.H., Prieto V.G.,
 RA Shea C.R., Berg J.N., Nevin N.C., Simpson S.A., Pasyk K.A.,
 RA Speer M.C., Peters K.G., Marchuk D.A.;
 RT "Allelic and locus heterogeneity in inherited venous malformations.";
 RL Hum. Mol. Genet. 8:1279-1289(1999).
 CC -1- FUNCTION: THIS PROTEIN IS A PROTEIN TYROSINE-KINASE TRANSMEMBRANE
 CC MAMMALIAN ENDOTHELIAL CELL LINEAGE MARKER. PROBABLY REGULATES
 CC ENDOTHELIAL CELL PROLIFERATION, DIFFERENTIATION AND GUIDES THE
 CC PROPER PATTERNING OF ENDOTHELIAL CELLS DURING BLOOD VESSEL
 CC FORMATION.
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein
 CC tyrosine phosphate.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN ENDOTHELIAL CELLS
 CC AND THEIR PROGENITORS, THE ANGIOBLASTS. HAS BEEN DIRECTLY FOUND
 CC IN PLACENTA AND LUNG, WITH A LOWER LEVEL IN UMBILICAL VEIN
 CC ENDOTHELIAL CELLS, BRAIN AND KIDNEY.
 CC -1- DISEASE: DEFECTS IN TEK ARE A CAUSE OF DOMINANTLY INHERITED VENOUS
 CC MALFORMATIONS (VMC1), AN ERROR OF VASCULAR MORPHOGENESIS
 CC CHARACTERIZED BY DILATED, SERPIGINOUS CHANNELS.
 CC -1- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
 CC DOMAIN.
 CC -1- SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 3 FIBRONECTIN TYPE III-LIKE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 3 EGF-LIKE DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: L06139; AAA61139.1; -
 DR HSSP: P11362; LFGL.
 DR MIM: 600221; -
 DR MIM: 600195; -
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR003961; FN_III.
 DR InterPro: IPR001245; Tyr_pkinase.
 DR Pfam: PF00008; EGF; 1.

DR Pfam: PF00041; fn3; 3.
 DR Pfam: PF00069; pkinase; 1.
 DR PRINTS: PR00109; TYRKINASE.
 DR SMART: SM00181; EGF; 2.
 DR SMART: SM00001; EGF_Like; 1.
 DR SMART: SM00060; FN3; 3.
 DR SMART: SM00219; TyRK; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00022; EGF_1; 3.
 DR PROSITE: PS01186; EGF_2; 3.
 KW Receptor; Tyrosine-protein kinase; Transferase; Signal; ATP-binding;
 KW Repeat; EGF-like domain; Transmembrane; Immunoglobulin domain;
 KW Glycoprotein; Phosphorylation; Multigene family; Disease mutation.
 FT SIGNAL 1 18
 FT CHAIN 19 1124
 FT DOMAIN 19 745
 FT TRANSMEM 746 770
 FT DOMAIN 771 1124
 FT DOMAIN 44 102
 FT DOMAIN 210 252
 FT DOMAIN 254 299
 FT DOMAIN 301 341
 FT DOMAIN 444 536
 FT DOMAIN 541 634
 FT DOMAIN 638 732
 FT DOMAIN 824 1096
 FT NP_BIND 830 838
 FT BINDING 855 855
 FT ACT_SITE 964 964
 FT CARBOHYD 140 140
 FT CARBOHYD 158 158
 FT CARBOHYD 339 339
 FT CARBOHYD 438 438
 FT CARBOHYD 464 464
 FT CARBOHYD 560 560
 FT CARBOHYD 596 596
 FT CARBOHYD 649 649
 FT CARBOHYD 691 691
 FT MOD_RES 992 992
 FT VARIANT 849 849
 FT VARIANT 897 897
 FT VARIANT /FTID=VAR_006352.
 FT VARIANT Y -> S (IN VMC1; ACTIVATING EFFECT).
 FT VARIANT /FTID=VAR_008716.
 SQ SEQUENCE 1124 AA; 125810 MW; 65BC05D18FAACCEC QRC64;
 Query Match 17.3%; Score 70; DB 1; Length 1124;
 Best Local Similarity 26.4%; Pred. No. 4;
 Matches 24; Conservative 4; Mismatches 37; Indels 26; Gaps 3;
 QY 3 EGYWDP-----LLGTCM-----SCKTCN-HQSORTCAAFCCR 36
 DB 212 EAQKWGPECNHLCTACNMNVGCHDTEGICPFPGMGRTEKACEHTEFGTKERCSCG 271
 QY 37 EQGKFVDHLRDCISCASCIGQHPKOCAYFC 67
 DB 272 EGCKSVFCLPDPYGCSCATGWGLQCNEAC 302
 RESULT 7
 ID ITBL_MOUSE STANDARD; PRT; 798 AA.
 AC P09055;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DE Integrin beta-1 precursor (Fibronectin receptor beta subunit)
 DE (CD29) (Integrin VLA-4 beta subunit).
 GN ITGB1.
 OS Mus musculus (Mouse).

ID PCK5_HUMAN STANDARD; PRT; 913 AA.
 AC Q92824; Q13527;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Proprotein convertase subtilisin/kexin type 5 precursor (EC 3.4.21.-)
 DE (Proprotein convertase PC5) (Subtilisin/kexin-like protease PC5)
 DE (Convertase PC5) (PC6) (hPC6).
 GN PCS5 OR PC5 OR PC6.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=T-cell;
 RX MEDLINE=96353880; PubMed=8755538;
 RA Miranda L., Wolf J., Pichuanes S., Duke R., Franzusoff A.;
 RT "Isolation of the human PC6 gene encoding the putative host protease
 for HIV-1 gp160 processing in CD4+ T lymphocytes.";
 RL Proc. Natl. Acad. Sci. U.S.A. 93:7695-7700(1996).
 RN [2]
 RP REVISIONS.
 RA Franzusoff A., Miranda L., Wolf J., Pichuanes S., Lu Y., Duke R.;
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 15-913 FROM N.A.
 RA Reudelhuber T.L.;
 RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: LIKELY TO REPRESENT A WIDESPREAD ENDOPEPTASE ACTIVITY
 CC WITHIN THE CONSTITUTIVE AND REGULATED SECRETORY PATHWAY. CAPABLE
 CC OF CLEAVAGE AT THE RX(K/R)R CONSENSUS MOTIF.
 CC -1- CATALYTIC ACTIVITY: RELEASE OF MATURE PROTEINS FROM THEIR
 CC PROTEINAINS BY CLEAVAGE OF ARG-XAA-YAA-ARG-1-ZAA BONDS, WHERE XAA
 CC CAN BE ANY AMINO ACID AND YAA IS ARG OR LYS.
 CC -1- SUBCELLULAR LOCATION: Secreted (By similarity).
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS SEEMS TO BE PRODUCED BY
 CC ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN T-LYMPHOCYTES.
 CC -1- DOMAIN: THE PROPEPTIDE DOMAIN ACTS AS AN INTRAMOLECULAR CHAPERONE
 CC ASSISTING THE FOLDING OF THE ZMOGEN WITHIN THE ENDOPLASMIC
 CC RETICULUM.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8; ALSO KNOWN AS THE
 CC SUBTILASE FAMILY.
 CC -1- SIMILARITY: CONTAINS 1 HOMO B/P DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U56387; AAC50643.2; -;
 DR EMBL; U49114; AAA91807.1; -;
 DR HSSP; Q99405; 1MPT.
 DR MIM; 600488; -;
 DR InterPro; IPR002174; Furin-like.
 DR InterPro; IPR002884; P_domain.
 DR InterPro; IPR000209; Peptidase_S8.
 DR Pfam; PF01483; P; 1.
 DR Pfam; PF00082; Peptidase_S8; 1.
 DR PRINTS; PR00723; SUBTILISIN.
 DR ProDom; PD000717; P_domain; 1.
 DR SMART; SM00261; FU; 5.
 DR PROSITE; PS00136; SUBTILASE_ASP; 1.
 DR PROSITE; PS00137; SUBTILASE_HIS; 1.
 DR PROSITE; PS00138; SUBTILASE_SER; 1.
 DR Hydrolase; Serine protease; Glycoprotein; Zymogen; Signal;
 KW Cleavage on pair of basic residues; Repeat.
 FT SIGNAL 1 32 BY SIMILARITY.
 FT PROPEP 33 114 BY SIMILARITY.

FT	CHAIN	115	913	PROTEIN CONVERTASE TYPE 5.	SUBTILISIN/KEXIN
FT	DOMAIN	115	454	CATALYTIC.	
FT	DOMAIN	462	600	HOMO B.	
FT	DOMAIN	636	913	CYS-RICH MOTIF (CRM)	REGION.
FT	SITE	114	115	CLEAVAGE (AUTO-) (BY SIMILARITY).	
FT	SITE	519	521	CELL ATTACHMENT SITE (POTENTIAL).	
FT	ACT_SITE	171	171	CHARGE RELAY SYSTEM (BY SIMILARITY).	
FT	ACT_SITE	212	212	CHARGE RELAY SYSTEM (BY SIMILARITY).	
FT	ACT_SITE	386	386	CHARGE RELAY SYSTEM (BY SIMILARITY).	
FT	CARBOHYD	225	225	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	381	381	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	665	665	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	752	752	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	802	802	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	852	852	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CONFLICT	118	118	S -> F (IN REF. 3).	
FT	CONFLICT	121	121	V -> A (IN REF. 3).	
FT	CONFLICT	511	511	R -> A (IN REF. 3).	
FT	CONFLICT	601	601	R -> Q (IN REF. 3).	
SQ	SEQUENCE	913 AA;	101775 MW;	21389264CAD7546C	CRC64;

Query Match 17.0%; Score 69; DB 1; Length 913;
 Best Local Similarity 26.0%; Pred. No. 4.3;
 Matches 19; Conservative 7; Mismatches 23; Indels 24; Gaps 5;

QY	1	CPEEQYWDPLLTGCMSCKTICNHSQRTCAAF--C--CRK-----	-----EQGFYD 43
Db	724	CPGSGYDPTKKNLCRCSENC-----KCTEFHNTCTCRDGLSLQSSRCVSCEGDGRFN	778

QY 44 HLLRDCIGCSASIC 56
 Db 779 G--QDCOPCHRC 789

RESULT 9
 TRA2_MOUSE STANDARD; PRT; 501 AA.
 AC P39429;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE TNF receptor associated factor 2 (TRAF2).
 GN TRAF2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RX MEDLINE=94349371; PubMed=8069916;
 RA Rothe M., Wong S.C., Henzel W.J., Goeddel D.V.;
 RT "A novel family of putative signal transducers associated with the
 RT cytoplasmic domain of the 75 kDa tumor necrosis factor receptor.";
 RL Cell 78:681-692(1994).
 CC -1- FUNCTION: SIGNAL TRANSDUCER ASSOCIATED WITH THE CYTOPLASMIC DOMAIN
 CC OF THE 75 KDA TUMOR NECROSIS FACTOR RECEPTOR (TNF-R2) AND
 CC ACTIVATES NF-KAPPA-B.
 CC -1- SUBUNIT: HOMODIMER OR HETERODIMER OF TRAF1 AND TRAF2. THIS
 CC HETEROCOMPLEX CAN BIND TO THE N-TERMINAL OF INHIBITOR OF APOPTOSIS
 CC PROTEINS 1 AND 2 (IAPs), TO RECRUIT THEM TO THE TUMOR NECROSIS
 CC FACTOR RECEPTOR 2 (TNFR2).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
 CC -1- SIMILARITY: CONTAINS A MATH/TRAF DOMAIN.
 CC -1- SIMILARITY: CONTAINS 2 TRAF-TYPE ZINC FINGERS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

RESULT 11
 LMAL_MOUSE STANDARD; PRT; 3084 AA.
 AC P19137;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Laminin alpha-1 chain precursor (Laminin A chain).
 GN LMAL OR LAMA-1 OR LAMA.
 OS Mus musculus (Mouse)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RN MEDLINE=89034134; PubMed=3182802;
 RA Sasaki M., Kleinman H.K., Huber H., Deutzmann R., Yamada Y.;
 RT "Laminin, a multidomain protein. The A chain has a unique globular
 domain and homology with the basement membrane proteoglycan and the
 laminin B chains."
 RT Laminin B chains."
 RL J. Biol. Chem. 263:16536-16544(1988).
 RN [2]
 RN SEQUENCE OF 1-339 FROM N.A.
 RX MEDLINE=88225080; PubMed=3267223;
 RA Hartl L., Oberbaumer I., Deutzmann R.;
 RT "The N terminus of laminin A chain is homologous to the B chains."
 RL Eur. J. Biochem. 173:629-635(1988).
 RN [3]
 RN SEQUENCE OF 2538-3084 FROM N.A., AND PARTIAL SEQUENCE.
 RX MEDLINE=89030693; PubMed=3181157;
 RA Deutzmann R., Huber J., Schmetz K.A., Oberbaumer I., Hartl L.;
 RT "Structural study of long arm fragments of laminin. Evidence for
 repetitive C-terminal sequences in the A-chain, not present in the B-
 chains."
 RL Eur. J. Biochem. 177:35-45(1988).
 CC -1- FUNCTION: BINDING TO CELLS VIA A HIGH AFFINITY RECEPTOR, LAMININ
 IS THOUGHT TO MEDIATE THE ATTACHMENT, MIGRATION, & ORGANIZATION OF
 CELLS INTO TISSUES DURING EMBRYONIC DEVELOPMENT BY INTERACTING
 WITH OTHER EXTRACELLULAR MATRIX COMPONENTS.
 CC -1- SUBUNIT: LAMININ IS A COMPLEX GLYCOPROTEIN, CONSISTING OF THREE
 DIFFERENT POLYPEPTIDE CHAINS (ALPHA, BETA, GAMMA), WHICH ARE BOUND
 TO EACH OTHER BY DISULFIDE BONDS INTO A CROSS-SHAPED MOLECULE
 COMPRISING ONE LONG & THREE SHORT ARMS WITH GLOBULES AT EACH END.
 CC THE ALPHA-1 CHAIN IS A SUBUNIT OF LAMININ-1 (EHS LAMININ) AND
 LAMININ-3 (S-LAMININ).
 CC -1- SUBCELLULAR LOCATION: Extracellular.
 CC -1- TISSUE SPECIFICITY: FOUND IN THE BASEMENT MEMBRANES (MAJOR
 COMPONENT).
 CC -1- DOMAIN: THE ALPHA-HELICAL DOMAINS I AND II ARE THOUGHT TO INTERACT
 WITH OTHER LAMININ CHAINS TO FORM A COILED COIL STRUCTURE.
 CC -1- DOMAIN: DOMAINS VI, IV AND G ARE GLOBULAR.
 CC -1- SIMILARITY: CONTAINS 1 LAMININ N-TERMINAL DOMAIN (DOMAIN VI).
 CC -1- SIMILARITY: CONTAINS 17 LAMININ EGF-LIKE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 2 LAMININ DOMAINS IV.
 CC -1- SIMILARITY: CONTAINS 5 LAMININ G-LIKE DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; J04064; AAA39410.1; -;
 DR EMBL; X07737; CAA30561.1; -;
 DR EMBL; X13459; CAA31807.1; -;
 DR EMBL; M36775; AAA39406.1; -;
 DR PIR; A31771; MMSA.
 DR HSSP; P02468; ITLE.
 DR MGD; MGI:99892; Lmal.
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR001886; LamNT.
 DR InterPro; IPR000034; Laminin_B.
 DR InterPro; IPR002049; Laminin_EGF.
 DR InterPro; IPR001791; Laminin_G.
 DR Pfam; PF00052; laminin_B; 2.
 DR Pfam; PF00053; laminin_EGF; 15.
 DR Pfam; PF00054; laminin_G; 5.
 DR Pfam; PF00055; laminin_Nterm; 1.
 DR PRINTS; PR00011; EGF_LAMININ.
 DR ProDom; PD002082; LamNT; 1.
 DR ProDom; PD003031; Laminin_B; 2.
 DR SMART; SM00180; EGF_Lam; 14.
 DR SMART; SM00281; LamB; 2.
 DR SMART; SM00282; LamG; 5.
 DR SMART; SM00136; LamNT; 1.
 DR PROSITE; PS00022; EGF_1; 11.
 DR PROSITE; PS01186; EGF_2; 3.
 DR PROSITE; PS01248; LAMININ_TYPE_EGF; 15.
 DR PROSITE; PS00025; LAM_G_DOMAIN; 5.
 KW Glycoprotein; Basement membrane; Extracellular matrix; Coiled coil;
 KW Laminin EGF-like domain; Cell adhesion; Repeat; Signal.
 FT SIGNAL 1 24 POTENTIAL.
 FT CHAIN 25 3084 LAMININ ALPHA-1 CHAIN
 FT MOD_RES 25 25 BLOCKED.
 FT DOMAIN 25 276 LAMININ N-TERMINAL (DOMAIN VI).
 FT DOMAIN 277 333 LAMININ EGF-LIKE 1.
 FT DOMAIN 334 403 LAMININ EGF-LIKE 2.
 FT DOMAIN 404 460 LAMININ EGF-LIKE 3.
 FT DOMAIN 461 509 LAMININ EGF-LIKE 4.
 FT DOMAIN 510 519 LAMININ EGF-LIKE 5 (N-TERMINAL).
 FT DOMAIN 520 715 LAMININ DOMAIN IV 1 (DOMAIN IV B).
 FT DOMAIN 716 748 LAMININ EGF-LIKE 5 (C-TERMINAL).
 FT DOMAIN 749 797 LAMININ EGF-LIKE 6.
 FT DOMAIN 798 855 LAMININ EGF-LIKE 7.
 FT DOMAIN 856 908 LAMININ EGF-LIKE 8.
 FT DOMAIN 909 957 LAMININ EGF-LIKE 9.
 FT DOMAIN 958 1004 LAMININ EGF-LIKE 10.
 FT DOMAIN 1005 1050 LAMININ EGF-LIKE 11.
 FT DOMAIN 1051 1096 LAMININ EGF-LIKE 12.
 FT DOMAIN 1097 1156 LAMININ EGF-LIKE 13.
 FT DOMAIN 1157 1166 LAMININ EGF-LIKE 14 (N-TERMINAL).
 FT DOMAIN 1167 1368 LAMININ DOMAIN IV 2 (DOMAIN IV A).
 FT DOMAIN 1369 1409 LAMININ EGF-LIKE 14 (C-TERMINAL).
 FT DOMAIN 1410 1458 LAMININ EGF-LIKE 15.
 FT DOMAIN 1459 1515 LAMININ EGF-LIKE 16.
 FT DOMAIN 1516 1562 LAMININ EGF-LIKE 17.
 FT DOMAIN 1564 2124 DOMAIN II AND I.
 FT DOMAIN 2125 2305 LAMININ G-LIKE 1.
 FT DOMAIN 2313 2489 LAMININ G-LIKE 2.
 FT DOMAIN 2494 2680 LAMININ G-LIKE 3.
 FT DOMAIN 2722 2894 LAMININ G-LIKE 4.
 FT DOMAIN 2899 3079 LAMININ G-LIKE 5.
 FT DOMAIN 1612 1820 COILED COIL (POTENTIAL).
 FT DOMAIN 1869 1903 COILED COIL (POTENTIAL).
 FT DOMAIN 2096 2128 COILED COIL (POTENTIAL).
 FT SITE 1147 1149 CELL ATTACHMENT SITE.
 FT DISULFID 277 286 BY SIMILARITY.
 FT DISULFID 279 297 BY SIMILARITY.
 FT DISULFID 299 308 BY SIMILARITY.
 FT DISULFID 311 331 BY SIMILARITY.
 FT DISULFID 334 343 BY SIMILARITY.
 FT DISULFID 336 368 BY SIMILARITY.
 FT DISULFID 371 380 BY SIMILARITY.
 FT DISULFID 383 401 BY SIMILARITY.
 FT DISULFID 404 416 BY SIMILARITY.
 FT DISULFID 406 434 BY SIMILARITY.
 FT DISULFID 436 445 BY SIMILARITY.
 FT DISULFID 448 458 BY SIMILARITY.
 FT DISULFID 461 474 BY SIMILARITY.
 FT DISULFID 463 478 BY SIMILARITY.
 FT DISULFID 480 489 BY SIMILARITY.
 FT DISULFID 492 507 BY SIMILARITY.
 FT DISULFID 749 758 BY SIMILARITY.


```

DR EMBL; L06864; AA53029.1; -.
DR EMBL; Z12608; CAA78249.1; -.
DR HSPG; P11362; IFGK.
DR MGI; MGI:95294; Egfr.
DR InterPro; IPR000494; EGFR_L.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002174; Furin-like.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00757; Furin-like; 1.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF01030; Recep_L_domain; 2.
DR SMART; SM00261; FU_3.
DR SMART; SM00219; Tyrc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW Transmembrane; Glycoprotein; Receptor; Signal; Transferase;
KW Tyrosine-protein kinase; ATP-binding; Phosphorylation; Repeat.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 1210 EPIDERMAL GROWTH FACTOR RECEPTOR.
FT DOMAIN 25 647 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 648 670 POTENTIAL.
FT DOMAIN 671 1210 CYTOPLASMIC (POTENTIAL).
FT REPEAT 75 300 APPROXIMATE.
FT REPEAT 390 600 APPROXIMATE.
FT DOMAIN 1028 1071 SER-RICH.
FT DOMAIN 714 981 PROTEIN KINASE.
FT NP_BIND 720 728 ATP (BY SIMILARITY).
FT BINDING 747 747 ATP (BY SIMILARITY).
FT ACT_SITE 839 839 BY SIMILARITY.
FT MOD_RES 680 680 PHOSPHORYLATION (BY PKC) (BY SIMILARITY).
FT MOD_RES 1092 1092 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT MOD_RES 1110 1110 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT MOD_RES 1172 1172 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT MOD_RES 1197 1197 PHOSPHORYLATION (AUTO-, MAJOR SITE) (BY SIMILARITY).
FT CARBOHYD 128 128 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 175 175 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 196 196 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 352 352 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 413 413 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 444 444 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 528 528 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 568 568 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 603 603 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 623 623 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 19 19 C -> S (IN REF. 2).
FT CONFLICT 539 539 C -> W (IN REF. 5).
FT CONFLICT 991 991 L -> F (IN REF. 4).
FT CONFLICT 1116 1117 HP -> DR (IN REF. 6).
SQ SEQUENCE 1210 AA; 134853 MW; 690E20D46DF2D2F5 CRC64;

Query Match 16.5%; Score 67; DB 1; Length 1210;
Best Local Similarity 27.2%; Pred. No. 8.7;
Matches 22; Conservative 9; Mismatches 30; Indels 20; Gaps 5;

QY 1 CPEQYWDPLLGTCMCKTICNHSQRTCAAF--CRKEQKGFYDLRLDCISCASIC-- 56
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 510 CSSGCGWGPEDCVSCQ---NYSRGECVEKCNILSGEPFVEN--SECIQCHPCPLP 564
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 57 -----GQHPK---QCAFY 66
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 565 QAMNITCTGRGPDNCIOCAHY 585
| | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 13
LMA2_HUMAN
ID LMA2_HUMAN STANDARD; PRT; 3110 AA.
AC P24043; Q14736; Q93022;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)

```

```

DE Laminin alpha-2 chain precursor (Laminin M chain) (Merosin heavy
DE chain).
DE LAMA2 OR LAMM.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=94124633; PubMed=8294519;
RA Vuolteenaho R., Nissinen M., Sainio K., Byers M., Eddy R.,
RA Hirvonen H., Shows T.B., Sariola H., Engvall E., Tryggvason K.;
RT "Human laminin M chain (merosin): complete primary structure,
RT chromosomal assignment, and expression of the M and A chain in human
RT fetal tissues.";
RL J. Cell Biol. 124:381-394(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97066955; PubMed=8910357;
RA Zhang X., Vuolteenaho R., Tryggvason K.;
RT "Structure of the human laminin alpha2-chain gene (LAMA2), which is
RT affected in congenital muscular dystrophy.";
RL J. Biol. Chem. 271:27664-27669(1996).
RN [3]
RP SEQUENCE OF 1981-3110 FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Placenta;
RX MEDLINE=90238994; PubMed=2185464;
RA Ehrig K., Leivo I., Argraves W.S., Ruoslahti E., Engvall E.;
RT "Merosin, a tissue-specific basement membrane protein, is a
RT laminin-like protein.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:3264-3268(1990).
RN [4]
RP VARIANTS GLN-545; HIS-619; LEU-919; HIS-2586 AND LYS-2614.
RA Panicker S.G., Mendell J.T., Chen L., Feng B., Sahenk Z.,
RA Marzluf G.A., Amato A.A., Mendell J.R.;
RT "Novel single base polymorphisms and rare sequence variants in
RT the laminin 2-chain coding region detected by RNA/SSCP analysis.";
RL Hum. Mutat. 13:174-174(1999).
RN [5]
RP ERRATUM.
RA Panicker S.G., Mendell J.T., Chen L., Feng B., Sahenk Z.,
RA Marzluf G.A., Amato A.A., Mendell J.R.;
RL Hum. Mutat. 13:340-340(1999).
CC -!- FUNCTION: BINDING TO CELLS VIA A HIGH AFFINITY RECEPTOR, LAMININ
CC IS THOUGHT TO MEDIATE THE ATTACHMENT, MIGRATION, & ORGANIZATION OF
CC CELLS INTO TISSUES DURING EMBRYONIC DEVELOPMENT BY INTERACTING
CC WITH OTHER EXTRACELLULAR MATRIX COMPONENTS.
CC -!- SUBUNIT: LAMININ IS A COMPLEX GLYCOPROTEIN, CONSISTING OF THREE
CC DIFFERENT POLYPEPTIDE CHAINS (ALPHA, BETA, GAMMA), WHICH ARE BOUND
CC TO EACH OTHER BY DISULFIDE BONDS INTO A CROSS-SHAPED MOLECULE
CC COMPRISING ONE LONG & THREE SHORT ARMS WITH GLOBULES AT EACH END.
CC THE ALPHA-2 CHAIN IS A SUBUNIT OF LAMININ-2 (MEROSIN) AND LAMININ-
CC 4 (S-MEROSIN)
CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR; FOUND IN THE BASEMENT
CC MEMBRANES (MAJOR COMPONENT).
CC -!- TISSUE SPECIFICITY: PLACENTA, STRIATED MUSCLE, PERIPHERAL NERVE,
CC CARDIAC MUSCLE, PANCREAS, LUNG, SPLEEN, KIDNEY, ADRENAL GLAND,
CC SKIN, TESTIS, MENINGES, CHOROID PLEXUS, AND SOME OTHER REGIONS OF
CC THE BRAIN; NOT IN LIVER, THYMUS AND BONE.
CC -!- DOMAIN: THE ALPHA-HELICAL DOMAINS I AND II ARE THOUGHT TO INTERACT
CC WITH OTHER LAMININ CHAINS TO FORM A COILED COIL STRUCTURE.
CC -!- DOMAIN: DOMAINS VI, IV AND G ARE GLOBULAR.
CC -!- DISEASE: DEFECTS IN LAMA2 ARE THE CAUSE OF MEROSIN-DEFICIENT
CC CONGENITAL MUSCULAR DYSTROPHY (MCMD).
CC -!- SIMILARITY: CONTAINS 1 LAMININ N-TERMINAL DOMAIN (DOMAIN VI).
CC -!- SIMILARITY: CONTAINS 17 LAMININ EGF-LIKE DOMAINS.
CC -!- SIMILARITY: CONTAINS 2 LAMININ DOMAINS IV.
CC -!- SIMILARITY: CONTAINS 5 LAMININ G-LIKE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```


Receptor; Proto-oncogene; Transmembrane; Chromosomal translocation;
Polymorphism.
TRANSMEM 55 77 POTENTIAL.
SITE 3 4 BREAKPOINT FOR TRANSLOCATION TO FORM
INTERLEUKIN 2/BCM ONCOGENE.
VARIANT 153 153 A -> T.
SEQUENCE 184 AA: 20138 MW; 277AF11E276D932 CRC64;

Query Match 16.4%; Score 66.5; DB 1; Length 184;
Best Local Similarity 32.4%; Pred. No. 2.2;
Matches 11; Conservative 8; Mismatches 14; Indels 1; Gaps 1;

OY 1 CPQEQWDPDLLGTCSCKTICNHQS-QRTCAAFK 33
I : I L I I I I I : I : I : I I I I
DBB 8 CSQNEFYDSILHACIPQLRCSNTPPLTCQRYC 41

RESULT 15
ITBL_FELCA
ID ITBL_FELCA STANDARD; PRT; 798 AA.
AC P53713;
DT 01-OCT-1996 (Rel. 34, Created)
FT 01-OCT-1996 (Rel. 34, Last sequence update)
DE 01-MAR-2002 (Rel. 41, Last annotation update)
DE Integrin beta-1 precursor (Fibronectin receptor beta subunit) (CD29) (Integrin VLA-4 beta subunit).
GN ITGB1.
OS Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
NCBI_TaxID=9685;
RN [1]
RP SEQUENCE FROM N.A.
RA Willett B.J.;
RL Submitted (JUL-1995) to the EMBL/GenBank/DDBJ databases.

-! FUNCTION: INTEGRINS ALPHA-1/BETA-1, ALPHA-2/BETA-1, ALPHA-10/BETA-1 AND ALPHA-11/BETA-1 ARE RECEPTORS FOR COLLAGEN. INTEGRINS ALPHA-1/BETA-1 AND ALPHA-2/BETA-2 RECOGNIZE THE PROLINE-HYDROXYLATED PEPTIDE G-F-P-G-E-R IN COLLAGEN. INTEGRINS ALPHA-2/BETA-1, ALPHA-3/BETA-1, ALPHA-4/BETA-1, ALPHA-5/BETA-1, ALPHA-8/BETA-1, ALPHA-10/BETA-1, ALPHA-11/BETA-1 AND ALPHA-V/BETA-1 ARE RECEPTORS FOR FIBRONECTIN. ALPHA-4/BETA-1 RECOGNIZES ONE OR MORE DOMAINS WITHIN THE ALTERNATIVELY SPLICED CS-1 AND CS-5 REGIONS OF FIBRONECTIN. INTEGRINS ALPHA-5/BETA-1 IS A RECEPTOR FOR FIBRINOGEN. INTEGRIN ALPHA-1/BETA-1, ALPHA-2/BETA-1, ALPHA-6/BETA-1 AND ALPHA-7/BETA-1 ARE RECEPTORS FOR LAMININ. INTEGRIN ALPHA-4/BETA-1 IS A RECEPTOR FOR VCAM1 AND RECOGNIZES THE SEQUENCE Q-I-D-S IN VCAM1. INTEGRIN ALPHA-9/BETA-1 IS A RECEPTOR FOR VCAM1, CYTOTACTIN AND OSTEOPONTIN. IT RECOGNIZES THE SEQUENCE A-E-I-D-G-I-E-L IN CYTOTACTIN. INTEGRIN ALPHA-3/BETA-1 IS A RECEPTOR FOR EPILGRIN AND THROMBOSPONDIN. INTEGRIN ALPHA-V/BETA-1 IS A RECEPTOR FOR VITRONECTIN. BETA-1 INTEGRINS RECOGNIZE THE SEQUENCE R-G-D IN A WIDE ARRAY OF LIGANDS.

-! SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. BETA-1 ASSOCIATES WITH EITHER ALPHA-1, ALPHA-2, ALPHA-3, ALPHA-4, ALPHA-5, ALPHA-6, ALPHA-7, ALPHA-8, ALPHA-9, ALPHA-10, ALPHA-11 OR ALPHA-V.

-! SUBCELLULAR LOCATION: Type I membrane protein.

-! SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.

-! SIMILARITY: CONTAINS 1 VWFA-LIKE DOMAIN.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its contents is in no way modified and this statement is not removed. Usage by individuals or commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to licensed@isb-sib.ch).

EMBL; U27351; AAC19407.1; -.
DR InterPro: IPR000561; EGF-like.

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 25, 2002, 16:24:30 ; Search time 89.98 Seconds
(without alignments)
128.814 Million cell updates/sec

Title: US-09-854-864-16
Perfect score: 405
Sequence: 1 CPEQYWDPLLTGTCMSCKTI.....DCISCASICGHPKQCAFC 67

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues 562222
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : SPTREMBL_19.*
- 1: sp_archaea.*
 - 2: sp_bacteria.*
 - 3: sp_fungi.*
 - 4: sp_human.*
 - 5: sp_invertebrate.*
 - 6: sp_mammal.*
 - 7: sp_mhc.*
 - 8: sp_organelle.*
 - 9: sp_phase.*
 - 10: sp_plant.*
 - 11: sp_rodent.*
 - 12: sp_virus.*
 - 13: sp_vertebrate.*
 - 14: sp_unclassified.*
 - 15: sp_rvirus.*
 - 16: sp_bacteriaph.*
 - 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	393	97.0	293	4	O14836
2	256	63.2	249	11	Q9DBZ3
3	256	63.2	249	11	Q9ET35
4	86.5	21.4	971	11	Q9Z0J1
5	84.5	20.9	971	4	Q95980
6	81.5	20.1	1299	5	Q26489
7	80	19.8	108	16	O07571
8	77	19.0	2664	5	Q26033
9	76.5	18.9	783	5	P92163
10	76	18.6	297	4	Q9HAV5
11	75.5	18.6	704	3	O74567
12	75	18.5	1069	5	Q9BPS2
13	74	18.3	723	11	Q9D2H5
14	71.5	17.7	213	11	Q99LE4
15	71.5	17.7	296	11	O35171
16	71.5	17.7	932	11	Q62030

17	71.5	17.7	1679	5	Q24301
18	71	17.5	210	4	Q9BYR0
19	71	17.5	761	6	Q95LQ2
20	71	17.5	1023	4	Q9ULI7
21	71	17.5	1704	5	Q94446
22	70.5	17.4	1362	13	Q9PVZ4
23	69.5	17.2	98	5	Q16939
24	69.5	17.2	109	2	O68643
25	69.5	17.2	166	4	Q9BYR3
26	69.5	17.2	271	5	Q9VU12
27	69	17.0	154	4	Q9BYP9
28	69	17.0	913	4	Q9SEP4
29	69	17.0	1137	4	Q9H8C1
30	69	17.0	1918	4	Q9BQM7
31	69	17.0	1925	4	Q9P2E3
32	68.5	16.9	438	11	Q922M5
33	68.5	16.9	508	11	O54896
34	68	16.8	433	11	Q912M6
35	67.5	16.7	146	10	Q40579
36	67.5	16.7	777	10	Q9LU038
37	67.5	16.7	2233	5	Q94711
38	67	16.5	136	11	Q99PK1
39	67	16.5	643	11	Q9ERV6
40	67	16.5	655	11	Q9WVF5
41	67	16.5	1083	11	Q9QW24
42	67	16.5	1210	11	Q9EP98
43	67	16.5	2108	13	Q98UI9
44	66.5	16.4	270	10	Q9LH57
45	66.5	16.4	282	10	Q9C7C1

ALIGNMENTS

RESULT 1
O14836 ID O14836 PRELIMINARY; PRT; 293 AA.
AC O14836;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE TRANSMEMBRANE ACTIVATOR AND CAML INTERACTOR.
GN TACI.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97458245; PubMed=9311921;
RA von Bulow G.U., Bram R.J.;
RT "NF-AT activation induced by a CAML-interacting member of the tumor
RT necrosis factor receptor superfamily.";
RL Science 278:138-141(1997);
DR EMBL; AF023614; AAC51790.1; --
DR InterPro; IPR001368; TNFR_C6.
DR PROSITE; P500652; TNFR_NGFR_1; UNKNOWN_1.
SQ SEQUENCE 293 AA; 31816 MW; 411799F3DE17A5EB CRC64;

Query Match 97.0%; Score 393; DB 4; Length 293;
Best Local Similarity 94.4%; Pred. No. 3e-44;
Matches 67; Conservative 0; Mismatches 0; Indels 4; Gaps 1;

QY	1	CPEQYWDPLLTGTCMSCKTI	CNHSORTCAFC	-----CRKEQKGYDHLRLDCISCASIC	56
Db	34	CPEQYWDPLLTGTCMSCKTI	CNHSORTCAFC	RSLSCKREQKGYDHLRLDCISCASIC	93
QY	57	GOHPKQCAFC	67		
Db	94	GOHPKQCAFC	104		

Q24301	drosophila
Q9BYR0	homo sapien
Q95LQ2	macaca fasc
Q9ULI7	homo sapien
Q94446	chironomus
Q9PVZ4	xenopus lae
Q16939	ancyllostoma
O68643	pseudomonas
Q9BYR3	homo sapien
Q9VU12	drosophila
Q9BYP9	homo sapien
Q9SEP4	homo sapien
Q9H8C1	homo sapien
Q9BQM7	homo sapien
Q9P2E3	homo sapien
Q922M5	mus musculus
O54896	mus musculus
Q912M6	rattus norv
Q40579	nicotiana t
Q9LU038	arabidopsis
Q94711	paramedius
Q99PK1	mus musculus
Q9ERV6	mus musculus
Q9WVF5	mus musculus
Q9QW24	rattus sp.
Q9EP98	mus musculus
Q98UI9	gallus gall
Q9LH57	arabidopsis
Q9C7C1	arabidopsis

```

RESULT 2
Q9DBZ3 ID Q9DBZ3 PRELIMINARY; PRT; 249 AA.
AC Q9DBZ3;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE 1200090E08RIK PROTEIN.
GN TNFRSF13B OR 1200090E08RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=LUNG;
RC MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaudo I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staabli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK004668; BAB23457.1; -.
DR MGD; MGI:1889411; Tnfrsf13b.
SQ SEQUENCE 249 AA; 26981 MW; 6F4290D719FEA037 CRC64;

Query Match 63.2%; Score 256; DB 11; Length 249;
Best Local Similarity 57.7%; Pred. No. 3.9e-26;
Matches 41; Conservative 11; Mismatches 15; Indels 4; Gaps 1;

QY 1 CPPEQYWDPLGLTGMCKTICNHQSQRCAAFQ----CRKEQGKFYDHLRLDCISCASIC 56
||:||||| :||| | :||||| || |||||:||||| |:|||
6 CPKQDYWDSSRKSCVSCALTCQRSQRCTCTDFCKFNCRKQGRYDHLGLGACVSCDSTC 65
||:||||| :||| | :||||| || |||||:||||| |:|||

QY 57 GQHPKQCAFC 67
||:|||||
DB 66 TQHPQCAHFC 76

RESULT 3
Q9ET35 ID Q9ET35 PRELIMINARY; PRT; 249 AA.
AC Q9ET35;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE TACI PROTEIN.
GN TNFRSF13B.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21177254; PubMed=10881172;
RA Yan M., Marsters S.A., Grewal I.S., Wang H., Ashkenazi A., Dixit V.M.;
RT "Identification of a receptor for Blys demonstrates a crucial role in

Query Match 63.2%; Score 256; DB 11; Length 249;
Best Local Similarity 57.7%; Pred. No. 3.9e-26;
Matches 41; Conservative 11; Mismatches 15; Indels 4; Gaps 1;

QY 1 CPPEQYWDPLGLTGMCKTICNHQSQRCAAFQ----CRKEQGKFYDHLRLDCISCASIC 56
||:||||| :||| | :||||| || |||||:||||| |:|||
6 CPKQDYWDSSRKSCVSCALTCQRSQRCTCTDFCKFNCRKQGRYDHLGLGACVSCDSTC 65
||:||||| :||| | :||||| || |||||:||||| |:|||

QY 57 GQHPKQCAFC 67
||:|||||
DB 66 TQHPQCAHFC 76

Query Match 21.4%; Score 86.5; DB 11; Length 971;
Best Local Similarity 26.8%; Pred. No. 0.0046;
Matches 26; Conservative 8; Mismatches 28; Indels 35; Gaps 6;

QY 1 CPPE--QYWDPL-----LGTG-----MSCKTICNHQSQRCAAFQCRKE 37
||:||||| :||| | :||||| || |||||:||||| |:|||
74 CPETMVEIWSMNSSLPVPFKKSDGWLGCCELAIGLECRQACKQASSKNDISKVCRKE 133
||:||||| :||| | :||||| || |||||:||||| |:|||

QY 38 QGKFYDHLRLDCIS-----CASICQGHKQCAFC 67
||:||||| :||| | :||||| || |||||:||||| |:|||
DB 134 ----YENALFSCISRNEMSGVCCSYAGHH-TNCREFC 165

RESULT 5
Q95980 ID Q95980 PRELIMINARY; PRT; 971 AA.
AC Q95980;
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)

```

```

RT humoral immunity.";
RL Nat. Immunol. 1:37-41(2000).
DR EMBL; AF257673; AAG0081.1; -.
DR MGD; MGI:1889411; Tnfrsf13b.
SQ SEQUENCE 249 AA; 26947 MW; CB2F2D61C2931D81 CRC64;

Query Match 63.2%; Score 256; DB 11; Length 249;
Best Local Similarity 57.7%; Pred. No. 3.9e-26;
Matches 41; Conservative 11; Mismatches 15; Indels 4; Gaps 1;

QY 1 CPPEQYWDPLGLTGMCKTICNHQSQRCAAFQ----CRKEQGKFYDHLRLDCISCASIC 56
||:||||| :||| | :||||| || |||||:||||| |:|||
6 CPKQDYWDSSRKSCVSCALTCQRSQRCTCTDFCKFNCRKQGRYDHLGLGACVSCDSTC 65
||:||||| :||| | :||||| || |||||:||||| |:|||

QY 57 GQHPKQCAFC 67
||:|||||
DB 66 TQHPQCAHFC 76

RESULT 4
Q9Z0J1 ID Q9Z0J1 PRELIMINARY; PRT; 971 AA.
AC Q9Z0J1;
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE MRECK PROTEIN PRECURSOR.
GN RECK OR MRECK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99007295; PubMed=9789069;
RA Takahashi C., Sheng Z., Horan T.P., Kitayama H., Maki M., Hitomi K.,
RA Kitaura Y., Takai S., Sasahara R.M., Horimoto A., Ikawa Y.,
RA Ratzkin B.J., Arakawa T., Noda M.;
RT "Regulation of matrix metalloproteinase-9 and inhibition of tumor
invasion by the membrane-anchored glycoprotein RECK.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:13221-13226(1998).
DR EMBL; AB006960; BAA34061.1; -.
DR MGD; MGI:1855698; Reck.
DR InterPro; IPR002350; kazal.
DR Pfam; PF00050; kazal; 2.
DR SMART; SM00280; KAZAL; 2.
DR PROSITE; PS00282; KAZAL; UNKNOWN_1.
KW Signal.
FT SIGNAL 1 26 POTENTIAL.
SQ SEQUENCE 971 AA; 106134 MW; 2FC8EBE38A20F86D CRC64;

Query Match 21.4%; Score 86.5; DB 11; Length 971;
Best Local Similarity 26.8%; Pred. No. 0.0046;
Matches 26; Conservative 8; Mismatches 28; Indels 35; Gaps 6;

QY 1 CPPE--QYWDPL-----LGTG-----MSCKTICNHQSQRCAAFQCRKE 37
||:||||| :||| | :||||| || |||||:||||| |:|||
74 CPETMVEIWSMNSSLPVPFKKSDGWLGCCELAIGLECRQACKQASSKNDISKVCRKE 133
||:||||| :||| | :||||| || |||||:||||| |:|||

QY 38 QGKFYDHLRLDCIS-----CASICQGHKQCAFC 67
||:||||| :||| | :||||| || |||||:||||| |:|||
DB 134 ----YENALFSCISRNEMSGVCCSYAGHH-TNCREFC 165

RESULT 5
Q95980 ID Q95980 PRELIMINARY; PRT; 971 AA.
AC Q95980;
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)

```

```
DE ST15 PROTEIN PRECURSOR.
GN ST15.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC
RX TISSUE=LUNG;
RA Takahashi C., Sheng Z., Horan T.P., Kitayama H., Maki M., Hitomi K.,
RA Kitaura Y., Takai S., Sasahara R.M., Horimoto A., Ikawa Y.,
RA Ratzkin B.J., Arakawa T., Noda M.;
RT "Regulation of matrix metalloproteinase-9 and inhibition of tumor
RL invasion by the membrane-anchored glycoprotein RECK.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:13221-13226(1998).
DR EMBL: D50406; BAA34060.1;
DR InterPro: IPR002350; kazal.
DR SMART: SM00280; KAZAL; 3.
DR PROSITE: PS00282; KAZAL; UNKNOWN_1.
KW SIGNAL.
FT SIGNAL.
SQ SEQUENCE 971 AA; 106457 MW; 173D47D6AE6F834 CRC64;

Query Match 20.9%; Score 84.5; DB 4; Length 971;
Best Local Similarity 25.8%; Pred. No. 0.0085;
Matches 25; Conservative 10; Mismatches 27; Indels 35; Gaps 6;

QY 1 CPPE-QYWDPL-----LGTC-----MSCKTICNHSQRTCAAFCKRKE 37
DB 74 CPPEQYWDPLGTCMSCKTICNHSQRTCAAFCKRKE 37
QY 38 QGRFYDHLRDCIS-----CASICGQHPKQCAAYFC 67
DB 134 ---YENALFSCISRNEMSGVCCSYAGHH-TNCREYC 165

RESULT 6
ID Q26489 PRELIMINARY; PRT; 1299 AA.
AC Q26489;
DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DE 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
DE ENDOPEPTIDASE FURIN.
GN FURIN
OS Spodoptera frugiperda (Fall armyworm).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Noctuoidea; Noctuidae; Amphipyrinae; Spodoptera.
OX NCBI_TaxID=7108;
RN [1]
RP SEQUENCE FROM N.A.
RC
RX TISSUE=SPIN.
RA Clepik M., Klenk H.;
RT "Cloning and functional characterization of FURIN from Spodoptera
RT frugiperda (Sf9) cells.";
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: Z68888; CAA93116.1;
DR HSP; Q99405; IMPT.
DR MEROPS: S08-UPB; -.
DR InterPro: IPR002174; Furin-like.
DR InterPro: IPR002029; Peptidase_S8.
DR InterPro: IPR002884; P_domain.
DR Pfam: PF01483; P; 1.
DR Pfam: PF00082; Peptidase_S8; 1.
DR PRINTS: PR00723; SUBTILISIN.
DR ProDom: PD000717; P_domain; 1.
DR SMART: SM00261; FU; 10.
DR PROSITE: PS00136; SUBTILASE_ASP; 1.
DR PROSITE: PS00137; SUBTILASE_HIS; 1.
DR PROSITE: PS00138; SUBTILASE_SER; 1.
```

```
KW Protease.
SQ SEQUENCE 1299 AA; 142020 MW; 4C3799C7B8C572AB CRC64;

Query Match 20.1%; Score 81.5; DB 5; Length 1299;
Best Local Similarity 30.1%; Pred. No. 0.028;
Matches 22; Conservative 5; Mismatches 33; Indels 13; Gaps 3;

QY 1 CPPEQYWDPLGTCMSCKTICNHSQRTCAAFCKRKEQGRFYDHLRDCISCAIC 56
DB 1116 CLGSQYDARTSGTCRCDASG-----RTCSGPGQFSCSTCSRPLDIHLRNNQCVCCSER 1170
QY 57 G----QHPKQCAAY 65
DB 1171 GVTNSTPTDCCCH 1183

RESULT 7
ID 007571 PRELIMINARY; PRT; 108 AA.
AC 007571;
DT 01-JUL-1997 (TEMBLrel. 04, Created)
DT 01-JUL-1997 (TEMBLrel. 04, Last sequence update)
DT 01-NOV-1998 (TEMBLrel. 08, Last annotation update)
DE HYPOTHETICAL 11.8 KDA PROTEIN.
GN YHJQ.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RA Noback M.A., Terpstra P., Holsappel S., Venema G., Bron S.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moser I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borriass R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Conneron I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA Guiseppe G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Maestl D., Nakai S., Noback M.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nagai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scroffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemura K.,
RA Takeuchi M., Takakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzsaecker T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the gram-positive bacterium Bacillus
RT subtilis.";
RL Nature 390:249-256(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
```

RA Kunst F., Ogasawara N., Yoshikawa H., Danchin A.;
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Y14081; CAA74479.1; -;
 DR EMBL; Z99109; CAB12900.1; -;
 KW Hypothetical protein: Complete proteome.
 SQ SEQUENCE 108 AA; 11839 MW; 78637689ABB0FD8F CRC64;

Query Match 19.8%; Score 80; DB 16; Length 108;
 Best Local Similarity 22.6%; Pred. No. 0.004;
 Matches 24; Conservative 7; Mismatches 25; Indels 50; Gaps 4;

OY 4 EGYWPLGLTGMCSKTIONH-----QSORTCAAF----- 33

DB 2 EGYSEACIEACIDCMKACNCFKLEESVQHHLSCGIRLDRECADICALAVKAMQTDSP 61

OY 34 -----CRKEQGF-YDHLRLDCISASICGQHPKQC 63

DB 62 FMKETCALCADICEACGTECGKHDH-----CQACAKACFTCAEQC 103

RESULT 8

O26033 PRELIMINARY; PRT; 2664 AA.

AC Q26033;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE VARIANT-SPECIFIC SURFACE PROTEIN.
 GN VAR-2.

OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

OX NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=FCR3;

RX MEDLINE=95330813; PubMed=7606788;

RA Su X.Z., Heatwole V.M., Wertheimer S.P., Guinet F., Herrfeldt J.A.,
 RA Peterson D.S., Ravetch J.A., Wellem's T.B.;

RT "The large diverse gene family var encodes proteins involved in
 RT cytoadherence and antigenic variation of Plasmodium falciparum-
 RT infected erythrocytes.;"

RL Cell 82:89-100(1995).;

DR EMBL; L40609; AAA75398.1; -;

DR InterPro; IPR004258; PFEMP.

DR Pfam; PF03011; PFEMP; 2.

SQ SEQUENCE 2664 AA; 302410 MW; 6EA2468511703091 CRC64;

Query Match 19.0%; Score 77; DB 5; Length 2664;
 Best Local Similarity 44.7%; Pred. No. 0.23;
 Matches 17; Conservative 4; Mismatches 9; Indels 8; Gaps 2;

OY 32 FCCRKEQGFYDHLRLDCISAS-----ICGQHPKQC 63

DB 1031 FC--KEQSLRELLURDCGCTTGKCNNDKCAKCDKQC 1066

RESULT 9

P92163 PRELIMINARY; PRT; 783 AA.

ID P92163;
 AC P92163;
 DT 01-MAY-1997 (TREMBlrel. 03, Created)
 DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE INTEGRIN BETA G SUBUNIT.

OS Strongylocentrotus purpuratus (Purple sea urchin).

OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;

OC Echinoidea; Euechinoidea; Echinacea; Echinoidea; Strongylocentrotidae;

OC Strongylocentrotus.

OX NCBI_TaxID=7668;

RN [1]
 RP SEQUENCE FROM N.A.

RA Marsden M., Burke R.D.;
 RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
 CC -!- PTM: THE CYSTEINE RESIDUES ARE INVOLVED IN INTRACHAIN DISULFIDE
 CC BONDS (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.

DR EMBL; U77584; AAB39739.1; -;
 DR EMBL; U77587; AAB39741.1; -;
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR002369; Integrin_B.
 DR InterPro; IPR001169; Integrin_beta_C.
 DR InterPro; IPR003659; PSI.

DR InterPro; IPR002035; VWFA.

DR Pfam; PF00362; Integrin_B; 1.

DR ProDom; PD001811; Integrin_B; 1.

DR SMART; SM00001; EGF-like; 1.

DR SMART; SM00187; INB; 1.

DR SMART; SM00423; PSI; 1.

DR SMART; SM00327; VW; 1.

DR PROSITE; PS00022; EGF_1; UNKNOWN_2.

DR PROSITE; PS01186; EGF_2; UNKNOWN_1.

DR PROSITE; PS00243; INTEGRIN_BETA_2.

KW Cell adhesion; Cytoskeleton; Extracellular matrix; Glycoprotein;

KW Integrin; Repeat; Transmembrane.

SQ SEQUENCE 783 AA; 85530 MW; BB045C6F6D8FBFB2 CRC64;

Query Match 18.9%; Score 76.5; DB 5; Length 783;
 Best Local Similarity 33.3%; Pred. No. 0.08;
 Matches 19; Conservative 9; Mismatches 24; Indels 5; Gaps 3;

OY 13 TCMCKTICNHQSORTCAAFCCRKEQGFY-DHLRLDCISASICGQHPK--QCA YF 66

DB 592 TCMGTDTICNGEGVICGE--CKCNAGSSYRGALCQDCPTCSGGQCSRNEECVQCKAF 646

RESULT 10

O9HAV5

ID O9HAV5 PRELIMINARY; PRT; 297 AA.

AC O9HAV5;

DT 01-MAR-2001 (TREMBlrel. 16, Created)

DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)

DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)

DE X-LINKED ECTODYSPLASIN-A2 RECEPTOR.

GN XEDAR.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]
 RP SEQUENCE FROM N.A.

RX MEDLINE=20495245; PubMed=11039935;

RA Yan M., Wang L.C., Hymowitz S.G., Schilbach S., Lee J., Goddard A.,

RA de Vos A.M., Gao W.Q., Dixit V.M.;

RT "Two-amino acid molecular switch in an epithelial morphogen that

RT regulates binding to two distinct receptors.;"

RL Science 290:523-527(2000).

DR EMBL; AF298812; AAG28761.1; -;

DR InterPro; IPR001368; TNFR_c6.

DR Pfam; PF00020; TNFR_c6; 2.

DR SMART; SM00208; TNFR; 2.

DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.

DR PROSITE; PS00050; TNFR_NGFR_2; 2.

KW Receptor.

SQ SEQUENCE 297 AA; 32728 MW; 0E71127C6C48240C CRC64;

Query Match 18.8%; Score 76; DB 4; Length 297;
 Best Local Similarity 29.9%; Pred. No. 0.036;
 Matches 20; Conservative 12; Mismatches 21; Indels 14; Gaps 6;

OY 1 CPPEQYWDPLLTGTCMSCKTIC--NHQSORTCA-----AFC--C--RKEQGFYDHLRLD 48

RN [1]
 RP SEQUENCE FROM N.A.

DR	SMART: SM00180; EGF_Lam; 8.
DR	SMART: SM00001; EGF_Like; 1.
DR	PROSITE; PS00022; EGF_1; UNKNOWN_6.
DR	PROSITE; PS01186; EGF_2; 1.
DR	PROSITE; PS01248; LAMININ_TYPE_EGF; UNKNOWN_8.
KW	EGF-like domain; Glycoprotein.
FT	NON_TER 1
SQ	SEQUENCE 1069 AA; 118994 MW; DF40F37A0CCA79AA CRC64;
<hr/>	
Query Match 18.5%; Score 75; DB 5; Length 1069;	
Best Local Similarity 29.7%; Pred.No. 0.17;	
Matches 19; Conservative 5; Mismatches 28; Indels 12; Gaps	
<hr/>	
Qy	1 CPEQYWDPLPGTCMSCKTICNHSQRTCAAFCCRKEQGKFYDHLRDCISCA-SICGQH 59 : : : 257 CADNYGDPLRGTCERK--CNENIDITKPGNC-----DPYTKGLCLQLYNTAGEH 305
Dd	
Qy	60 PKQC 63
Dd	
Dd	306 CDVC 309
<hr/>	
RESULT 13	
Q9D2H5	PRELIMINARY; PRT; 723 AA.
ID Q9D2H5	
AC Q9D2H5;	
DT 01-JUN-2001 (TEMBUREl. 17, Created)	
DT 01-JUN-2001 (TEMBUREl. 17, Last sequence update)	
DT 01-DEC-2001 (TEMBUREl. 17, Last annotation update)	
DE 4930486B16RIK PROTEIN.	
GN 4930486B16RIK.	
DN 4930486B16RIK.	
OS Mus musculus (Mouse).	
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
ON NCBI_TaxID=10090;	
RX [1]	
RP SEQUENCE FROM N.A.	
RC SFPAIN-C57BL/6J; TISSUE=TESTIS;	
RX MEDLINE=11085660; PubMed=11217851;	
RA Kawai J., Shingagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,	
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,	
RA Aizawa K., Ikawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,	
RA Saito T., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R.,	
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,	
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiya H.,	
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,	
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,	
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,	
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,	
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,	
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,	
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,	
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,	
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,	
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,	
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji M., Kohtsuki S.,	
RA Hayashizaki Y.;	
"Functional annotation of a full-length mouse cDNA collection.";	
Nature 409:685-690(2001).	
RL EMBL; AK019654; BAB31824.1; -	
DR MGP; MG1:1926161; 4930486B16RIK.	
DR InterPro; IPR003961; FN.III	
DR InterPro; IPR000315; znf_box.	
DR Pfam; PF00041; fn3; 1.	
DR Pfam; PF00643; zf-B_box; 2.	
DR SMART; SM00336; BBOX; 2.	
DR SMART; SM0060; FN3; 1.	
DR SMART; SM00184; RING; 2.	
DR PROSITE; PS00518; ZINC_FINGER_C3HC4; 1.	
SW zinc-finger.	
KO SEQUENCE 723 AA; 82966 MW; OF6BGCD0191A6D222 CRD64;	

```

Query Match      18.3%; Score 74; DB 11; Length 723;
Best Local Similarity 32.2%; Pred. No. 0.16;
Matches 19; Conservative 8; Mismatches 22; Indels 10; Gaps 4;

QY      9  PLLGTCMSCKTICNHQSQRTCAATCCRKEQGFYDHLRLDQCISCAISGQHPKQCA YFC 67
      | | : | | | : | | | : | | : | | : | | : | |
Db      19  PRLFSLCLCKCFI--FTSERNTCTCFP-----PYKDE--RNCQFCHTCAENP-NCHWCC 67

RESULT      14
Q99LE4
ID      Q99LE4      PRELIMINARY;      PRT;      213 AA.
AC      Q99LE4;
DT      01-JUN-2001 (TReMBLrel. 17, Created)
DT      01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT      01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE      SIMILAR TO SUBTILISIN-LIKE ENDOPROTEASE (FRAGMENT) .
OS      Mus musculus (Mouse) .
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

```

```

RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
DR ENBL; BC003302; AAH03302.1; -.
DR MEROPS; S08.075; -.
DR InterPro; IPR000345; CytC_heme_bind.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR002174; Furin-like.
DR InterPro; IPR000822; ZnF-C2H2.
DR SMART; SM00181; EGF; 3.
DR SMART; SM00261; FU; 4.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
FW Protease.
KT NON_TER
SQ SEQUENCE 213 AA; 23781 MW; 39ALD5235454ECB6 CRC64;

Query Match 17.7%; Score 71.5; DB 11; Length 213;
Best Local Similarity 29.6%; Pred. No. 0.1;
Matches 16; Conservative 7; Mismatches 24; Indels 7;

QY 14 CMSCKT-ICNHQSORTCAAFCCREQGKPYDHLLRDCTISC-----ASTCGQHPKQC 63
      |||: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8 CLSCRGGFYHHOENTNTCVTLCK---PAGLYADESORLCILRCHPCSKCKVCDEPKC 63

```

RESULT	15
035171	
ID	O35171 PRELIMINARY; PRT; 296 AA.
AC	O35171;
DT	01-JAN-1998 (TReMBRel. 05, Created)
DT	01-JAN-1998 (TReMBRel. 05, Last sequence update)
DT	01-DEC-2001 (TReMBRel. 19, Last annotation update)
DE	PACE4 (FRAGMENT).
GN	PCSK6 OR PACE4.
OS	Mus musculus (Mouse).
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX	NCBI_TaxID=10090;
RP	[1]
RC	SEQUENCE FROM N.A.
RX	STRAIN=BALE/C.
RX	MEDLINE=98053888; PubMed=9393739;
RA	Hubbard F.C., Goodrow T.L., Liu S.C., Brilliant M.H., Basset P.,
RA	Maixs R.E., Klein-Szanto A.J.;
RT	"Expression of PACE4 in chemically induced carcinomas is associated
RT	with spindle cell tumor conversion and increased invasive ability.";
RL	Cancer Res. 57:5226-5231(1997).

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 25, 2002, 16:16:09 ; Search time 88.08 Seconds
(without alignments)

84.491 Million cell updates/sec

Title: US-09-854-864-16

Perfect score: 405

Sequence: 1 CPEEQYWDPLLTGCMCKT1.....DCISCASICGHPKQCAVFC 67

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_032802.*

```

1: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	393	97.0	166	19	Human lymphocyte s
2	393	97.0	265	22	Human TACI splice
3	393	97.0	293	19	Human lymphocyte s
4	393	97.0	293	21	Human neutrokin-a
5	393	97.0	293	21	A transmembrane ac
6	393	97.0	293	22	Human TACI protein
7	393	97.0	293	22	Human tumour necro
8	256	63.2	249	21	A murine tnfr4, a
9	204	50.4	247	21	Human BR43x2, an i
10	84.5	20.9	220	22	Human secreted pro
11	76.5	18.9	325	19	Homo sapiens vascu

12	76.5	18.9	325	22	AAV97572	Human VEGF-D prote
13	76.5	18.9	354	19	AAW49036	Homo sapiens vascu
14	76.5	18.9	354	19	AAW53241	Homo sapiens vascu
15	76.5	18.9	354	19	AAW44293	Human vascular end
16	76.5	18.9	354	21	AAW10849	Human VEGD protein
17	76.5	18.9	354	21	AAW29049	Human VEGF-D prote
18	76.5	18.9	354	21	AAW70750	Human prepro-vascu
19	76.5	18.9	354	21	AAW70983	Human vascular end
20	76.5	18.9	354	22	AAU08441	Polypeptide for hu
21	76.5	18.9	354	22	AAW70685	Human vascular end
22	76.5	18.9	354	22	AAW70750	Human VEGF-DI prot
23	76.5	18.9	354	22	AAW37606	Human VEGF-D. Hom
24	76.5	18.9	620	18	AAW14994	Human c-Fos induce
25	76	18.8	77	21	AAW77467	Human Rank-like pr
26	76	18.8	173	22	AAU03118	Composite protein
27	76	18.8	197	21	AAW01421	Human TANGO 140-2.
28	76	18.8	206	21	AAW01420	Human TANGO 140-1.
29	76	18.8	231	21	AAW77468	Human Rank-like pr
30	76	18.8	267	22	AAU03114	Human uterine myom
31	76	18.8	269	22	AAU03106	Human uterine myom
32	76	18.8	297	22	AAU03113	Human uterine myom
33	76	18.8	297	22	AAW29534	Human TNFR homolog
34	76	18.8	299	21	AAW30547	Amino acid sequenc
35	76	18.8	299	21	AAW33477	Human PRO5727 prot
36	76	18.8	299	22	AAU03116	Composite protein
37	76	18.8	299	22	AAW29533	Human TNFR homolog
38	74.5	18.4	969	14	AAW41662	Paired basic amino
39	72	17.8	100	22	ABW41017	Peptide #8523 enco
40	72	17.8	100	22	AAW34792	Peptide #8829 enco
41	71.5	17.7	1679	22	ABW60498	Drosophila melanog
42	71.5	17.7	1679	22	ABW60502	Drosophila melanog
43	71	17.5	1281	22	AAE10608	Human novel KTAAL2
44	71	17.5	1281	22	AAE10610	Human novel KTAAL2
45	71	17.5	1617	22	AAE07870	Novel human protei

ALIGNMENTS

RESULT	1
AAW75785	AAW75785 standard; Protein; 166 AA.
ID	AAW75785 standard; Protein; 166 AA.
XX	AAW75785;
AC	AAW75785;
XX	18-JAN-1999 (first entry)
DT	18-JAN-1999 (first entry)
XX	Human lymphocyte surface receptor extracellular domain.
DE	TACI; transmembrane activator and CAML-interactor;
XX	calcium signal-modulating cyclophilin ligand; human;
KW	lymphocyte surface receptor; human; B-cell; B lymphocyte;
KW	infection; cancer; rheumatoid arthritis; autoimmune disease;
KW	glomerulonephritis; immunosuppressive; graft versus host disease;
KW	transplant rejection; therapy; signal transduction.
XX	Homo sapiens.
OS	Homo sapiens.
XX	WO9839361-A1.
PN	WO9839361-A1.
XX	11-SEP-1998.
PD	11-SEP-1998.
XX	03-MAR-1998; 98WO-US04270.
PF	03-MAR-1998; 98WO-US04270.
XX	03-MAR-1997; 97US-0810572.
PR	03-MAR-1997; 97US-0810572.
XX	(SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
PA	(SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
XX	Bram RJ, Von Bulow G;
PI	Bram RJ, Von Bulow G;
XX	WPI; 1998-506346/43.
DR	WPI; 1998-506346/43.
XX	N-PSDB; AAV57330.
XX	N-PSDB; AAV57330.

PT New isolated transmembrane activator protein - used to develop
PT products for treating e.g. infections, cancers, autoimmune and
PT inflammatory conditions, transplant rejection or graft-versus-host
PT disease

PS Claim 8; Page 73; 89pp; English.

XX This is the amino acid sequence of the N-terminal, i.e. the
CC extracellular, domain of novel human transmembrane activator and
CC CAML-interactor (TACI) protein (see AAW75783). TACI is a lymphocyte
CC receptor protein that is involved in the calcium activation pathway.
CC It is normally present in B-lymphocytes, and to a much lesser extent
CC in immature T-lymphocytes, and can therefore be targeted to
CC specifically regulate B cell responses without affecting T cell
CC activity. The extracellular domain of TACI functions as a binding
CC site for a ligand that stimulates the activation of the cell by
CC inducing the binding of the C-terminal portion (see AAW75784) of
CC TACI to the N-terminal domain of CAML. A recombinant form of the
CC extracellular portion of TACI acts as a dominant-negative or
CC blocking agent and acts to suppress the immune system. It can be
CC used to treat or prevent autoimmune disease, graft rejection or
CC graft versus host disease. The extracellular region is also used
CC in a claimed method for identifying a ligand for TACI, in which
CC binding of a candidate molecule is determined by detecting cellular
CC activation of the AP-1, CAMP or NF-KB pathway, of NF-AT
CC transcription factor, or of NF-AT dependent transcription.

XX Sequence 166 AA;

Query Match 97.0%; Score 393; DB 19; Length 166;
Best Local Similarity 94.4%; Pred. No. 6.4e-33;
Matches 67; Conservative 0; Mismatches 0; Indels 4; Gaps 1;

Qy 1 CPEQYWDPLLTGTCMSCKTICNHQSQTCAAFPC-----CRKEQGKFYDHLRLDCISCASIC 56
|||||
Db 34 cpeeqywdplltgctmsckticnhqsrtaafcrslsrckegkfkydhlrlcdciscasic 93
|||||

Qy 57 GOHPKQCAAYFC 67
|||||
Db 94 gqhpqcaayfc 104

RESULT 2
AAE09244
ID AAE09244 standard; Protein; 265 AA.

XX AAE09244;
XX
XX 19-NOV-2001 (first entry)
XX Human TACI splice variant protein.
XX Human; TNF: tumour necrosis factor; TALL-1; APRIL; TNF receptor;
KW TNFR; TACI; BCMA; therapy; cancer; leukaemia; myeloma; lymphoma;
KW autoimmune disease; rheumatoid arthritis; multiple sclerosis;
KW psoriasis.

OS Homo sapiens.
XX WO200160397-A1.
XX 23-AUG-2001.
XX 28-NOV-2000; 2000WO-US32378.
XX 16-FEB-2000; 2000US-0182938.
PR 22-AUG-2000; 2000US-0226986.
XX (GFTH) GENENTECH INC.

XX Ashkenazi AJ, Dodge KH, Grewal I, Kim KJ, Marsters SA, Pitti RM;
PI Yan M;

XX WPI; 2001-541628/60.
DR Inhibiting or neutralizing TALL-1 or APRIL polypeptide biological
XX activity for treating autoimmune disorders and cancer, comprises
PT exposing the cells to TALL-1 or APRIL polypeptide agonists or
PT antagonists -
XX Example 1; Fig 6; 160pp; English.

PS The invention relates to methods of using one or more agonists or
XX antagonists to modulate the activity of the members of TNF (tumour
CC necrosis factor) especially TALL-1, APRIL and TNF receptor (TNFR)
CC e.g. TACI or BCMA. The method is useful for treating pathological
CC conditions or diseases associated with increased TALL-1 and APRIL
CC expression or activity. TALL-1 and APRIL antagonists are used to
CC block the interaction between APRIL and TALL-1 with TACI or BCMA.
CC They are useful for treating a mammal suffering from cancer such
CC as leukaemia, lymphoma, myeloma, cancers of lung and colon and
CC autoimmune diseases e.g. rheumatoid arthritis, multiple sclerosis,
CC psoriasis and lupus erythematosus. The present sequence is human
CC TACI splice variant protein.

XX Sequence 265 AA;

Query Match 97.0%; Score 393; DB 22; Length 265;
Best Local Similarity 94.4%; Pred. No. 9.9e-33;
Matches 67; Conservative 0; Mismatches 0; Indels 4; Gaps 1;

Qy 1 CPEQYWDPLLTGTCMSCKTICNHQSQTCAAFPC-----CRKEQGKFYDHLRLDCISCASIC 56
|||||
Db 34 cpeeqywdplltgctmsckticnhqsrtaafcrslsrckegkfkydhlrlcdciscasic 93
|||||

Qy 57 GOHPKQCAAYFC 67
|||||
Db 94 gqhpqcaayfc 104

RESULT 3
AAW75783
ID AAW75783 standard; Protein; 293 AA.

XX AAW75783;
XX
XX 18-JAN-1999 (first entry)
XX Human lymphocyte surface receptor TACI.
XX TACI; transmembrane activator and CAML-interactor;
KW calcium signal-modulating cyclophilin ligand; human;
KW lymphocyte surface receptor; human; B-cell; B lymphocyte;
KW infection; cancer; rheumatoid arthritis; autoimmune disease;
KW glomerulonephritis; immunosuppressive; graft versus host disease;
KW transplant rejection; therapy.

OS Homo sapiens.
XX
XX Key Location/Qualifiers
FH 1..166
FT Domain /label= Extracellular_domain
FT /note= "Claim 8"
FT Domain 167..186
FT /label= Transmembrane_domain
FT 187..294
FT /label= Cytoplasmic_domain
FT /note= "Claim 6"
FT Peptide 34..71
FT /note= "TNFR_NGFR motif"

XX WO9839361-A1.
XX 11-SEP-1998.
PD

XX PF 03-MAR-1998; 98WO-US04270.
 XX PR 03-MAR-1997; 97US-0810572.
 XX PA (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
 XX PI Bram RJ, Von Bulow G;
 XX DR WPI; 1998-506346/43.
 XX DR N-PSDB; AAV57328.
 XX PT New isolated transmembrane activator protein - used to develop
 PT products for treating e.g. infections, cancers, autoimmune and
 PT inflammatory conditions, transplant rejection or graft-versus-host
 PT disease
 XX PS Claim 20; Fig 2a; 89pp; English.
 XX CC This is the amino acid sequence of novel human transmembrane
 CC activator and CAML-interactor (TACI) protein, a lymphocyte receptor
 CC protein that is involved in the calcium activation pathway. TACI
 CC is normally present in B-lymphocytes, and to a much lesser extent
 CC in immature T-lymphocytes, and can therefore be targeted to
 CC specifically regulate B cell responses without affecting T cell
 CC activity. TACI cDNA (seev57328) was isolated from a B-lymphocyte
 CC cDNA library using a yeast two-hybrid assay. Also claimed are
 CC the C-terminal (see AAW75784) and N-terminal (see AAW75785) fragments
 CC of TACI, recombinant DNA constructs, unicellular hosts, and
 CC antibodies to TACI protein. Methods are claimed for identifying a
 CC ligand for TACI and for identifying immunosuppressive drugs that
 CC selectively block the action of B lymphocytes without affecting
 CC mature T lymphocytes. TACI can be activated to increase immune
 CC system activity, e.g. for treating infections or cancers. It can
 CC be blocked to provide immunosuppression, e.g. for treating
 CC autoimmune and inflammatory conditions such as immune complex-
 CC induced vasculitis, glomerulonephritis, haemolytic anaemia,
 CC myasthenia gravis, type II collagen-induced arthritis, experimental
 CC allergic and hyperacute xenograft rejection, rheumatoid arthritis,
 CC systemic lupus erythematosus, transplant rejection, cancer or
 CC graft versus host disease.
 XX CC Sequence 293 AA;
 SQ

Query Match 97.0%; Score 393; DB 19; Length 293;
 Best Local Similarity 94.4%; Pred. No. 1.1e-32;
 Matches 67; Conservative 0; Mismatches 0; Indels 4; Gaps 1;

QY 1 CPPEQYWDPLLGTCMCKTICNHQSORTCAAF- ---CRKEQGKFDYDHLRDCISCASIC 56
 |||||||||||||||||||||||||||||||||||
 Db 34 cpeegywdpllgtcmscktkcnhqsrtcaafcrs|scrkegkfydhllrdciscasic 93

QY 57 GQHPKQCAVFC 67
 |||||||||||
 Db 94 gqhpqcafc 104

RESULT 4
 AAB36312
 ID AAB36312 standard; Protein; 293 AA.
 XX AC AAB36312;
 XX DT 26-FEB-2001 (first entry)
 XX DE Human neutrokin-alpha binding protein TR17 SEQ ID NO:2.
 XX CC Human; neutrokin-alpha binding protein; NAR protein; TR17; cytostatic;
 KW immunosuppressive; nontropic; neuroprotective; antiviral; antiallergic;
 KW hepatotropic; antidiabetic; antiinflammatory; antiulcer; cardiant;
 KW ophthalmological; gene therapy; immunodeficiency disorder; diagnosis;
 KW autoimmune disorder.

XX OS Homo sapiens.
 XX PN WO200058362-A1.
 XX PD 05-OCT-2000.
 XX PF 24-MAR-2000; 2000WO-US07966.
 XX PR 26-MAR-1999; 99US-0126599.
 XX PR 10-MAR-2000; 2000US-0188208.
 XX PA (HUMA-) HUMAN GENOME SCI INC.
 XX PI Ruben SM, Ullrich S, Baker K;
 XX DR WPI: 2000-602359/57.
 XX DR N-PSDB; AAC64602.
 PT Nucleic acid encoding a neutrokin-alpha receptor (NAR) such as TR17,
 PT useful for producing TR17 protein which is used in the treatment and
 PT diagnosis of autoimmune and immunodeficiency disorders -
 XX Claim 1; Fig 1; 398pp; English.
 XX CC The present sequence represents the human neutrokin-alpha binding (NAR)
 CC protein designated TR17. TR17 has cytostatic, immunosuppressive,
 CC nontropic, neuroprotective, antiviral, antiallergic, hepatotropic,
 CC antidiabetic, antiinflammatory, antiulcer, cardiant and ophthalmological
 CC activities and can be used in gene therapy. The TR17 protein and
 CC antibodies are useful for treating and diagnosing immunodeficiency
 CC disorders and autoimmune disorders. The TR17 polypeptides,
 CC polynucleotides, antibodies, agonists and/or antagonists are used for
 CC treating various other diseases defined in the specification and as
 CC research tools for studying the phenotypic effects that result from
 CC inhibiting TR17/TR17 ligand interactions on various cell types.
 XX CC Sequence 293 AA;
 SQ

Query Match 97.0%; Score 393; DB 21; Length 293;
 Best Local Similarity 94.4%; Pred. No. 1.1e-32;
 Matches 67; Conservative 0; Mismatches 0; Indels 4; Gaps 1;

QY 1 CPPEQYWDPLLGTCMCKTICNHQSORTCAAF- ---CRKEQGKFDYDHLRDCISCASIC 56
 |||||||||||||||||||||||||||||||||||
 Db 34 cpeegywdpllgtcmscktkcnhqsrtcaafcrs|scrkegkfydhllrdciscasic 93

QY 57 GQHPKQCAVFC 67
 |||||||||||
 Db 94 gqhpqcafc 104

RESULT 5
 AAY94000
 ID AAY94000 standard; Protein; 293 AA.
 XX AC AAY94000;
 XX DT 20-OCT-2000 (first entry)
 XX DE A transmembrane activator and CAML-interactor (TACI).
 XX CC Human; BR43x2; TACI receptor; extracellular domain; BCMA; B cell protein;
 KW transmembrane activator and CAML-interactor; tumour necrosis factor; TNF;
 KW znfa activity; antibody production; autoimmune disease; amyloidosis;
 KW systemic lupus erythematosus; myasthenia gravis; multiple sclerosis;
 KW rheumatoid arthritis; asthma; bronchitis; emphysema; pyelonephritis;
 KW end stage renal failure; glomerulonephritis; vasculitis; nephritis;
 KW renal neoplasm; multiple myeloma; lymphoma; light chain neuropathy;
 KW immune response; immunosuppression; graft rejection; joint pain;
 KW graft versus host disease; inflammation; swelling; anaemia; septic shock;
 KW insulin dependent diabetes mellitus; Crohn's disease; hypertension;

KW renal artery stenosis; occlusion; cholesterol; renal emboli.
 XX Homo sapiens.
 XX WO200040716-A2.
 XX 13-JUL-2000.
 PD 07-JAN-2000; 2000WO-US00396.
 XX 07-JAN-1999; 99US-0226533.
 PR (ZYMO) ZYMOGENETICS INC.
 PA Gross JA, Xu W, Madden K, Yee DP;
 XX WPI; 2000-452538/39.
 DR N-PSDB; AAA58558.
 XX Inhibiting ztnf4 activity in a mammal, to treat autoimmune diseases,
 PT renal disease, graft versus host disease, and inflammation, comprises
 PT administering a BR43x2, TACI or BCMA extracellular domain polypeptide -
 XX Disclosure; Page 149-150; 175pp; English.
 XX The present sequence represents a human transmembrane activator and
 CC CAML-interactor (TACI) receptor. TACI is a tumour necrosis factor (TNF)
 CC receptor. The extracellular domains of BR43x2 (an isoform of TACI), TACI
 CC or BCMA (a related B cell protein) receptor contain a cysteine rich
 CC domain, and are used for inhibiting ztnf4 activity. Ztnf4 is a TNF
 CC ligand. They may also be used for inhibiting BR43x2, TACI or BCMA
 CC receptor-ligand engagement associated with activated or resting B
 CC lymphocytes, effector T-cells, or with antibody production. The
 CC antibody production is associated with an autoimmune disease selected
 CC from systemic lupus erythematosus, myasthenia gravis, multiple sclerosis
 CC and rheumatoid arthritis. The ztnf4 activity and BR43x2, TACI or BCMA
 CC receptor-ligand engagement is associated with asthma, bronchitis,
 CC nephritis, end stage renal failure, glomerulonephritis, vasculitis,
 CC light chain neuropathy, amyloidosis, multiple myelomas, lymphomas,
 CC immunosuppression, graft rejection, graft versus host disease,
 CC inflammation, insulin dependent diabetes mellitus, Crohn's disease, joint
 CC pain, swelling, anaemia, or septic shock. BR43x2, TACI, and BCMA
 CC polypeptides, fusions, antibodies, agonists or antagonists can be used
 CC to treat hypertension, renal artery stenosis, or occlusion, and
 CC cholesterol or renal emboli.

SQ Sequence 293 AA;

Query Match 97.0%; Score 393; DB 21; Length 293;
 Best Local Similarity 94.4%; Pred. No. 1.1e-32;
 Matches 67; Conservative 0; Mismatches 0; Indels 4; Gaps 1;

QY 1 CPBEQYWDPLLTGCMSCKTCNHQSQTCAAFCAFC---CRREQGKFYDHLRDCISCASIC 56
 |||||||
 Db 34 cpeeqywdplgtcmcsctkcnhqsrtcaafcrslscrkeqgkfydhllrdciscasic 93
 |||||||
 QY 57 GQHPKQCAAYFC 67
 |||||||
 Db 94 gqhpqkqayfc 104

RESULT 6
 AAE09240
 ID AAE09240 standard; Protein; 293 AA.
 XX
 AC AAE09240;

XX 19-NOV-2001 (first entry)
 DT Human TACI protein.
 XX

KW Human; TNF; tumour necrosis factor; TALL-1; APRIL; TNF receptor;
 KW TNFR; TACI; BCMA; therapy; cancer; leukaemia; myeloma; lymphoma;
 KW autoimmune disease; rheumatoid arthritis; multiple sclerosis;
 XX psoriasis.
 XX Homo sapiens.
 XX WO200160397-A1.
 PN 23-AUG-2001.
 PD 28-NOV-2000; 2000WO-US23278.
 XX 16-FEB-2000; 2000US-0182938.
 PR 22-AUG-2000; 2000US-0226986.
 XX (GETH) GENENTECH INC.
 PA Ashkenazi AJ, Dodge KH, Grewal I, Kim KJ, Marsters SA, Pitti RM;
 PI Yan M;
 XX WPI; 2001-541628/60.
 DR N-PSDB; AAD15901.
 XX Inhibiting or neutralizing TALL-1 or APRIL polypeptide biological
 PT activity, for treating autoimmune disorders and cancer, comprises
 PT exposing the cells to TALL-1 or APRIL polypeptide agonists or
 PT antagonists -
 XX Example 1; Fig 1; 160pp; English.

XX The invention relates to methods of using one or more agonists or
 CC antagonists to modulate the activity of the members of TNF (tumour
 CC necrosis factor) especially TALL-1, APRIL and TNF receptor (TNFR)
 CC e.g. TACI or BCMA. The method is useful for treating pathological
 CC conditions or diseases associated with increased TALL-1 and APRIL
 CC expression or activity. TALL-1 and APRIL antagonists are used to
 CC block the interaction between APRIL and TALL-1 with TACI or BCMA.
 CC They are useful for treating a mammal suffering from cancer such
 CC as leukaemia, lymphoma, myeloma, cancers of lung and colon and
 CC autoimmune diseases e.g. rheumatoid arthritis, multiple sclerosis,
 CC psoriasis and lupus erythematosus. The present sequence is human
 CC TACI protein.

SQ Sequence 293 AA;

Query Match 97.0%; Score 393; DB 22; Length 293;
 Best Local Similarity 94.4%; Pred. No. 1.1e-32;
 Matches 67; Conservative 0; Mismatches 0; Indels 4; Gaps 1;

QY 1 CPBEQYWDPLLTGCMSCKTCNHQSQTCAAFCAFC---CRREQGKFYDHLRDCISCASIC 56
 |||||||
 Db 34 cpeeqywdplgtcmcsctkcnhqsrtcaafcrslscrkeqgkfydhllrdciscasic 93
 |||||||
 QY 57 GQHPKQCAAYFC 67
 |||||||
 Db 94 gqhpqkqayfc 104

RESULT 7
 AAY71914
 ID AAY71914 standard; Protein; 293 AA.

XX AAY71914;

XX 26-MAR-2001 (first entry)

DT Human tumour necrosis factor receptor (TACI) protein.

XX Human; transmembrane activator and CAML interactor; TACI;
 KW tumour necrosis factor receptor; TNF; autoimmune disease; diabetes;
 KW calcium-signal modulating cyclophilin ligand; CAML; viral infection;

KW neutrokin alpha polypeptide; TACI-Ligand; TACI-L; cytostatic; therapy;
KW neuroprotective; antidiabetic; antiviral; antiinflammatory; tumour;
KW antiarthritic; antirheumatic; immunosuppressive; multiple sclerosis;
KW rheumatoid arthritis; graft rejection; inflammation; cell proliferation;
KW cell death; immunoglobulin E-mediated allergic reaction; IgE.

XX Homo sapiens.

XX Key Location/Qualifiers

FH Domain 2..166

FT /label= Extracellular domain

FT /note= "Binds with amino acids 123-285 of extracellular
FT domain of TACI-L"

XX WO200067034-A1.

PN 09-NOV-2000.

XX 14-APR-2000; 2000WO-US10282.

XX 30-APR-1999; 99US-0302863.

XX (IMMU) IMMUNEX CORP.

XX Goodwin RG, Din WS;

XX WPI; 2001-016005/02.

DR N-PSDB; AAD02006.

XX Use of new interactions between tumour necrosis factor receptors (TACI)
PT and TACI ligands to screen candidate molecules for determining agonist
PT and antagonist interactions which are used for treating inflammation -

XX Claim 10; Fig 1b; 46pp; English.

XX The present sequence is a human tumour necrosis factor receptor (TACI)
CC protein. TACI (Transmembrane activator and calcium-signal modulating
CC cyclophilin ligand (CAML)-interactor) forms a complex with neutrokin
CC alpha polypeptide (TACI-Ligand). The antagonist or agonist of
CC TACI/TACI-L complex is useful for modulating an intracellular signalling
CC cascade mediated by TACI/TACI-L complex. Antagonists of TACI/TACI-L
CC complex are used to inhibit the interaction between TACI and TACI-L for
CC therapeutic purposes to treat tumour and tumour metastasis and to combat
CC various autoimmune diseases e.g. multiple sclerosis and diabetes, as
CC well as other disorders, such as viral infection, rheumatoid arthritis,
CC graft rejection, and immunoglobulin (Ig) E-mediated allergic reactions,
CC and inflammation. The interaction is used to study cellular processes
CC associated with tumour necrosis factor (TNF)-receptors such as immune
CC regulation, cell proliferation, cell death and inflammatory responses.
CC The interaction between the extracellular region of TACI and TACI-L can
CC be used to further develop understanding of which cell types TACI-L
CC acts upon.

XX Sequence 293 AA;

Query Match 97.0%; Score 393; DB 22; Length 293;

Best Local Similarity 94.4%; Pred. No. 1.1e-32;

Matches 67; Conservative 0; Mismatches 0; Indels 4; Gaps 1;

QY 1 CPSEQYWDPLLGTCMSCKTICNHQSORTCAAFCAFC----CRKEQKQFYDHLRDCISASC 56

Db 34 cpeeqywdpllgtcmsckticnhqsrtaafcrsqrctdckfncrkeqgrydhllgacvsdsc 93

QY 57 GQHPKQCAyFC 67

Db 94 gqnpkqca yfc 104

RESULT 8

ID AAY94006

XX AAY94006 standard; Protein; 249 AA.

AC AAY94006;

XX 20-OCT-2000 (first entry)

XX A murine ztnf4, a tumour necrosis factor ligand.

XX Human; BR43x2; TACI receptor; extracellular domain; BCMA; B cell protein;
KW Transmembrane activator and CAML-interactor; tumour necrosis factor; TNF;
KW ztnf4 activity; antibody production; autoimmune disease; amyloidosis;
KW systemic lupus erythematosus; myasthenia gravis; multiple sclerosis;
KW rheumatoid arthritis; asthma; bronchitis; emphysema; pyelonephritis;
KW end stage renal failure; glomerulonephritis; vasculitis; neuropathy;
KW renal neoplasm; multiple myeloma; lymphoma; light chain neuropathy;
KW immune response; immunosuppression; graft rejection; joint pain;
KW graft versus host disease; inflammation; swelling; anaemia; septic shock;
KW insulin dependent diabetes mellitus; Crohn's disease; hypertension;
KW renal artery stenosis; occlusion; cholesterol; renal emboli.

XX Mus musculus.

XX WO200040716-A2.

XX 13-JUL-2000.

XX 07-JAN-2000; 2000WO-US00396.

XX 07-JAN-1999; 99US-0226533.

XX (ZYMO) ZYMOGENETICS INC.

XX Gross JA, Xu W, Madden K, Yee DP;

XX WPI; 2000-452538/39.

DR N-PSDB; AAA58566.

XX Inhibiting ztnf4 activity in a mammal, to treat autoimmune diseases,
PT renal disease, graft versus host disease, and inflammation, comprises
PT administering a BR43x2, TACI or BCMA extracellular domain polypeptide -
XX Disclosure; Page 163; 175pp; English.

XX The present sequence represents murine ztnf4, a tumour necrosis factor
CC ligand. The extracellular domains of BR43x2 (an isoform of the
CC transmembrane activator and CAML-interactor (TACI) receptor), TACI or
CC BCMA (a related B cell protein) contain a cysteine rich domain, and are
CC used for inhibiting ztnf4 activity. They may also be used for
CC inhibiting BR43x2, TACI or BCMA receptor-ligand engagement associated
CC with activated or resting B lymphocytes, effector T-cells, or with
CC antibody production. The antibody production is associated with an
CC autoimmune disease selected from systemic lupus erythematosus, myasthenia
CC gravis, multiple sclerosis and rheumatoid arthritis. The ztnf4 activity
CC and BR43x2, TACI or BCMA receptor-ligand engagement is associated with
CC asthma, bronchitis, emphysema, end stage renal failure,
CC glomerulonephritis, vasculitis, nephritis, pyelonephritis, renal
CC neoplasms, multiple myelomas, lymphomas, light chain neuropathy, graft
CC amyloidosis, moderating immune response, immunosuppression, graft
CC rejection, graft versus host disease, inflammation, insulin dependent
CC diabetes mellitus, Crohn's disease, joint pain, swelling, anaemia, or
CC septic shock. BR43x2, TACI, and BCMA polypeptides, fusions, antibodies,
CC agonists or antagonists can be used to treat hypertension, renal artery
CC stenosis, or occlusion, and cholesterol or renal emboli.

XX Sequence 249 AA;

Query Match 63.2%; Score 256; DB 21; Length 249;

Best Local Similarity 57.7%; Pred. No. 1e-18;

Matches 41; Conservative 11; Mismatches 15; Indels 4; Gaps 1;

QY 1 CPSEQYWDPLLGTCMSCKTICNHQSORTCAAFCAFC----CRKEQKQFYDHLRDCISASC 56

Db 6 cpkdywdsrkskcvscaltcsqrctdckfncrkeqgrydhllgacvsdsc 65

QY 57 GQHPKQCAAYFC 67
Db 66 tqhpqcahfc 76

RESULT 9
AAI93998
ID AAY93998 standard; Protein; 247 AA.
XX
XX AAY93998;
XX
XX 20-OCT-2000 (first entry)
XX
XX Human BR43x2, an isoform of the TACI receptor.
XX
XX Human: BR43x2; TACI receptor; extracellular domain; BCMA: B cell protein; transmembrane activator and CAML-interactor; tumour necrosis factor; TNF; ztnf4 activity; antibody production; autoimmune disease; amyloidosis; systemic lupus erythematosus; myasthenia gravis; multiple sclerosis; rheumatoid arthritis; asthma; bronchitis; emphysema; pyelonephritis; end stage renal failure; glomerulonephritis; vasculitis; nephritis; renal neoplasm; multiple myeloma; lymphoma; light chain neuropathy; immune response; immunosuppression; graft rejection; joint pain; graft versus host disease; inflammation; swelling; anaemia; septic shock; insulin dependent diabetes mellitus; Crohn's disease; hypertension; renal artery stenosis; occlusion; cholesterol; renal emboli.
XX
XX Homo sapiens.
XX
XX
XX Key Location/Qualifiers
XX Domain 1..120
XX Region /note= "extracellular domain"
XX Region 25..58
XX Domain /note= "cysteine-rich pseudo repeat"
XX Domain 121..133
XX Domain /note= "transmembrane domain"
XX Domain 134..247
XX Domain /note= "cytoplasmic domain"

WO200040716-A2.
XX
XX 13-JUL-2000.
XX
XX 07-JAN-2000; 2000WO-US00396.
XX
XX 07-JAN-1999; 99US-0226533.
XX
XX (ZYMO) ZYMOGENETICS INC.
XX
XX Gross JA, Xu W, Madden K, Yee DP;
XX
XX WPI; 2000-452538/39.
XX
XX N-PSDB; AAA58556.
XX
XX Inhibiting ztnf4 activity in a mammal, to treat autoimmune diseases, renal disease, graft versus host disease, and inflammation, comprises administering a BR43x2, TACI or BCMA extracellular domain polypeptide -
XX
XX Claim 62; Page 145; 175pp; English.

The present sequence represents a human BR43x2 polypeptide, which is an isoform of the transmembrane activator and CAML-interactor (TACI) receptor. TACI is a tumour necrosis factor (TNF) receptor. The extracellular domains of BR43x2, TACI or BCMA (a related B cell protein) contain a cysteine rich domain, and are used for inhibiting ztnf4 activity. ztnf4 is a TNF ligand. They may also be used for inhibiting BR43x2, TACI or BCMA receptor-ligand engagement associated with activated or resting B lymphocytes, effector T-cells, or with antibody production. The antibody production is associated with an autoimmune disease selected from systemic lupus erythematosus, myasthenia gravis, multiple sclerosis and rheumatoid arthritis. The ztnf4 activity and BR43x2, TACI or BCMA receptor-ligand engagement is associated with asthma, bronchitis, emphysema, end stage renal failure, glomerulonephritis, vasculitis,

CC nephritis, pyelonephritis, renal neoplasms, multiple myelomas, lymphomas, light chain neuropathy, amyloidosis, moderating immune response, immunosuppression, graft rejection, graft versus host disease, inflammation, insulin dependent diabetes mellitus, Crohn's disease, joint pain, swelling, anaemia, or septic shock. BR43x2, TACI, and BCMA polypeptides, fusions, antibodies, agonists or antagonists can be used to treat hypertension, renal artery stenosis, or occlusion, and cholesterol or renal emboli.
XX
XX Sequence 247 AA;
SQ

Query Match 50.4%; Score 204; DB 21; Length 247;
Best Local Similarity 100.0%; Pred. No. 2.1e-13;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 CRKEQGKFYDHLRDCISCASICGQHPKQCAAYFC 67
Db 25 crkeggkfydhllrdciscasacigqhpqcaayfc 58
|||||
|||||

RESULT 10
AAB65001
ID AAB65001 standard; protein; 220 AA.
XX
XX AAB65001;
XX
XX 23-MAR-2001 (first entry)
XX
XX Human secreted protein #9.
XX
XX Secreted protein; gene therapy; vaccine; cancer; leukemia; autoimmune disease; allergy; inflammation; graft rejection; hyperproliferation; cardiovascular; infection.
XX
XX Homo sapiens.
XX
XX WO200075375-A1.
XX
XX 14-DEC-2000.
XX
XX 02-JUN-2000; 2000WO-US15187.
XX
XX 07-JUN-1999; 99US-0137725.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Ruben SM, Birse CE, Duan RD, Soppet DR, Rosen CA, Shi Y; Lafleur DW, Olsen HS, Ebner R, Florence KA, Ni J, Young PE;
XX
XX WPI; 2001-061741/07.
XX
XX Nucleic acids encoding 26 human secreted polypeptides, useful e.g. for preventing, diagnosing and/or treating cancers and for promoting wound healing -
XX
XX Claim 1; Page 457; 530pp; English.

The present invention relates to 26 secreted human proteins. The proteins may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate polypeptide expression. For example, they may be used in gene therapy or in vaccines. Typical of diseases which are potentially treatable are cancers (including leukemia), autoimmune diseases, allergies, inflammation, graft rejection, hyperproliferation, cardiovascular diseases (particularly critical limb ischemia and coronary disease) and any involving abnormal angiogenesis, neurodegeneration and/or infectious diseases.
XX
XX Sequence 220 AA;
SQ

Query Match 20.9%; Score 84.5; DB 22; Length 220;

Best Local Similarity 25.8%; Pred. No. 0.33;
Matches 25; Conservative 10; Mismatches 27; Indels 35; Gaps 6;
QY 1 CPEE--QYWDPL-----LQTC-----MSCKTTCNHQSORTCAAFCCRKE 37
Db 74 cpetmveimcmasslpgvfkksdywvgicclatalecrqackqasskndiskvrke 133
QY 38 QGKFYDHLRDCIS-----CASICGQHPKQAYFC 67
Db 134 ----yenalfscirnmegsvccsyaghh-tncrey 165
RESULT 11
AAW53240
ID AAW53240 standard; Protein; 325 AA.
XX AC AAW53240;
XX DT 03-AUG-1998 (first entry)
XX DE Homo sapiens vascular endothelial growth factor D (VEGF-D).
XX KW vascular endothelial growth factor; VEGF-D; angiogenesis;
XX KW modification; acceleration; wound healing; tissue; organ;
XX KW transplants; collateral circulation; infarction; arterial stenosis;
XX KW coronary artery disease; inhibition; cancer; treatment;
XX KW diabetic retinopathy; lung disorders; blood circulation;
XX KW gaseous exchange; chronic obstructive airway disease;
XX KW intestinal malabsorptive syndrome; biopsy; metastatic risk;
XX KW detection; diagnosis; congestive heart failure.
XX OS Homo sapiens.
XX FH Key
XX FH Location/Qualifiers
XX FH 126..128
XX FH /note= "potential N-linked glycosylation site"
XX FT Region
XX FT 136..158
XX FT /note= "potential N-linked glycosylation site"
XX FT Region
XX FT 258..260
XX FT /note= "potential N-linked glycosylation site"
XX PN WO9807832-A1.
XX PD 26-FEB-1998.
XX PF 21-AUG-1997; 97WO-US14696.
XX PR 01-JUL-1997; 97US-0051426.
XX PR 23-AUG-1996; 96AU-0001825.
XX PR 23-AUG-1996; 96US-0023751.
XX PR 11-NOV-1996; 96AU-0003554.
XX PR 14-NOV-1996; 96US-0031097.
XX PR 05-FEB-1997; 97AU-0004954.
XX PR 10-FEB-1997; 97US-0038814.
XX PR 19-JUN-1997; 97AU-0007435.
XX PA {LUDW-} LUDWIG INST CANCER RES.
XX PA {UYHE-} UNIV HELSINKI LICENSING LTD.
XX PI Achen MG, Alitalo K, Stacker SA, Wilks AF;
XX DR WPI; 1998-179057/16.
XX DR N-PSDB; AAV20806.
XX PT New isolated vascular endothelial growth factor-D - used to develop
XX PT products for use in e.g. modifying angiogenesis or treating lung,
XX PT heart or intestinal disorders
XX PS Claim 16; Pages 57-58; 101pp; English.
XX CC The sequence is that of human breast vascular endothelial growth factor
XX CC D (VEGF-D). VEGF-D can be used for e.g. acceleration of angiogenesis
XX CC in wound healing, tissue or organ transplantation, or to establish

collateral circulation in tissue infarction or arterial stenosis,
CC such as coronary artery disease, and inhibition of angiogenesis in
CC the treatment of cancer or of diabetic retinopathy. It can also be
CC used in the treatment of lung disorders to improve blood circulation
CC in the lung and/or gaseous exchange between the lungs and the blood
CC stream or to improve blood circulation to the heart and O2 gas
CC permeability in cases of cardiac insufficiency, to improve blood
CC flow and gaseous exchange in chronic obstructive airway disease,
CC or to treat malabsorptive syndromes in the intestinal tract.
CC Quantitation of VEGF-D in cancer biopsy specimens may be useful
CC as an indicator of future metastatic risk. Antagonists can be used
CC for treating e.g. conditions such as congestive heart failure,
CC involving accumulations of fluid in the lung resulting from
CC increases in vascular permeability. The products can also be used
CC for detection and diagnosis.
XX SQ Sequence 325 AA;
Query Match 18.9%; Score 76.5; DB 19; Length 325;
Best Local Similarity 24.8%; Pred. No. 3.2;
Matches 28; Conservative 6; Mismatches 22; Indels 57; Gaps 7;
QY 1 CPEEQVWD-----PLIGT-----CMSCKTTCNH- 23
Db 193 cpidmldwsnkckvleqeenplagtedshlqepalcgphmmfdehrccecv-cktpcpkd 251
QY 24 --QSORTCAAF-----CCRKEQGFYDHL- -RDCISCASTCQHPKQCA 64
Db 252 lldhpknscfeckesletccqk-----hkifhpdtscedrcphtprca 297
RESULT 12
AAV97572
ID AAV97572 standard; Protein; 325 AA.
XX AC AAV97572;
XX DT 05-APR-2001 (first entry)
XX DE Human VEGF-D protein sequence.
XX KW Human; angiogenic protein; wound healing; vascular tissue repair;
XX KW peripheral arterial disease; critical limb ischaemia; coronary disease;
XX KW angiogenesis; tumour; inflammation; diabetic retinopathy; psoriasis;
XX KW rheumatoid arthritis; autoimmune disease; allergy; cancer; therapy;
XX KW infectious disease; neurodegeneration;
XX KW vascular endothelial growth factor-D; VEGF-D.
XX OS Homo sapiens.
XX PN WO2000075163-A1.
XX PD 14-DEC-2000.
XX PF 01-JUN-2000; 2000WO-US14925.
XX PR 03-JUN-1999; 99US-0137796.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI Rosen CA, Ruben SM, Hu J, Cao L;
XX DR WPI; 2001-071057/08.
XX DR N-PSDB; AAA91006.
XX PT New nucleic acid encoding angiogenic proteins, useful e.g. for
XX PT promoting healing of wounds and treating peripheral arterial disease,
XX PT critical limb ischaemia or coronary disease -
XX PS Claim 11; Page 226-227; 244pp; English.
XX CC This sequence is vascular endothelial growth factor-D (VEGF-D),

CC which is an angiogenic protein of the invention. The angiogenic proteins
CC and the DNA sequences encoding them, are used to prevent, treat or
CC ameliorate disease and to detect diseases, or susceptibility, by
CC detecting mutations or the presence or amount of angiogenic protein
CC expression. Particularly they are used to stimulate wound healing,
CC growth of damaged bone and tissue, and for repair of vascular tissue,
CC especially peripheral arterial disease, critical limb ischaemia or
CC coronary disease. Antagonists of the sequences are used to inhibit
CC angiogenesis in tumours and to treat inflammation (where associated with
CC increased vascular permeability), diabetic retinopathy, rheumatoid
CC arthritis or psoriasis. Agonists are also useful for stimulating
CC (lymph)angiogenesis. The proteins are also used to identify specific
CC binding agents (potential therapeutic agents) and to raise antibodies.
CC The antibodies are useful as therapeutic (antagonists; for detection,
CC purification and targeting of proteins for in vivo or in vitro diagnosis
CC (including imaging) or for therapy (including when linked to e.g. a label
CC or cytotoxin); and for immunotyping of cells, e.g. for detecting minimal
CC residual disease or haematopoietic progenitor/stem cells. It is also
CC contemplated that the sequences might be useful for treating a very wide
CC range of other disorders, e.g. autoimmune diseases; allergy; cancer;
CC infectious diseases (viral, bacterial, fungal or parasitic);
CC neurodegeneration, also as chemotactic agents or for stimulating
CC regeneration of the nervous system etc.
XX
SQ Sequence 325 AA;

Query Match 18.9%; Score 76.5; DB 22; Length 325;
Best Local Similarity 24.8%; Pred. No. 3.2;
Matches 28; Conservative 6; Mismatches 22; Indels 57; Gaps 7;
QY 1 CPEQYWD-----PLLGT-----CMSCRTICNH- 23
Db 193 cpdmlwdsnkckvqlqeenplagtdshlqepalcpghmmfdrcecv-cktpcpkd 251
QY 24 --QSORTCAAF-----CCRKEGKGYDHL--RDCISCAICGQHPKQCA 64
Db 252 liqhpknscfckesletccqk-----hklfhpdtcscedrcpfhtrpca 297

RESULT 13
AAW49036
ID AAW49036 standard; Protein; 354 AA.
XX
AC AAW49036;
XX
DT 26-OCT-1998 (first entry)
XX
DE Human zvegf2 growth factor.
XX
KW Human zvegf2 growth factor; mitogen; fibroblast; smooth muscle cell;
KW venous stasis ulcer; diabetic ulcer; skin wound; chemotactic effect;
KW angiogenic effect; tumour; diabetic retinopathy; psoriasis; arthritis;
KW scleroderma.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..23
FT Peptide /note= "Signal peptide"
FT Peptide 24..108
FT Peptide /note= "Pro-region"
FT Binding-site 109..197
FT Region /note= "Receptor binding domain"
FT Region 206..256
FT Region /note= "Cysteine-rich domain"
FT Region 257..274
FT Region /note= "Balbiani ring motif"
FT Region 275..294
FT Region /note= "Balbiani ring motif"
FT Region 295..354
FT Region /note= "Cysteine-rich domain"
XX

PN WO9824811-A2.
XX
PD 11-JUN-1998.
XX
PF 20-NOV-1997; 97WO-US20888.
XX
PR 18-SEP-1997; 97US-0933455.
PR 06-DEC-1996; 96US-0759657.
XX
PA (ZYMO) ZYMOGENETICS INC.
XX
PI Conklin DC, Gilbert T, Hart CE, Nygaard S, Sheppard PO;
XX WPI; 1998-333256/29.
DR N-PSDB; AAV32823.
XX
PT New isolated vascular endothelial growth factor - used to develop
PT products for treating e.g. wounds, burns, myocardial infarction,
PT tumours, psoriasis, arthritis, restenosis or organ transplants
XX Claim 1; Pages 53-54; 77pp; English.
CC The present sequence represents a human zvegf2 growth factor encoded
CC by the zvegf2 cDNA which was isolated from a human heart cDNA library.
CC zvegf2 protein in a dimeric form acts as a mitogen for fibroblasts or
CC smooth muscle cells. zvegf2 is claimed to be useful for stimulating the
CC revascularisation of tissue or the re-endothelialisation of vascular
CC tissue. zvegf2 is particularly claimed to be useful for the treatment
CC of full-thickness skin wounds, including venous stasis ulcers and
CC diabetic ulcers. The zvegf2 protein is also claimed to be useful as an
CC additive in tissue adhesives for promoting revascularisation of the
CC healing tissue. Antagonists against zvegf2 can be used to block its
CC mitogenic, chemotactic and angiogenic effects. The antagonists may
CC therefore be useful for reducing growth of solid tumours by inhibiting
CC neovascularisation of the developing tumour or by directly blocking
CC tumour cell growth, in the treatment of diabetic retinopathy, psoriasis,
CC arthritis, and scleroderma.
XX
SQ Sequence 354 AA;

Query Match 18.9%; Score 76.5; DB 19; Length 354;
Best Local Similarity 24.8%; Pred. No. 3.4;
Matches 28; Conservative 6; Mismatches 22; Indels 57; Gaps 7;
QY 1 CPEQYWD-----PLLGT-----CMSCRTICNH- 23
Db 222 cpdmlwdsnkckvqlqeenplagtdshlqepalcpghmmfdrcecv-cktpcpkd 280
QY 24 --QSORTCAAF-----CCRKEGKGYDHL--RDCISCAICGQHPKQCA 64
Db 281 liqhpknscfckesletccqk-----hklfhpdtcscedrcpfhtrpca 326
RESULT 14
AAW53241
ID AAW53241 standard; Protein; 354 AA.
XX
AC AAW53241;
XX
DT 03-AUG-1998 (first entry)
XX
DE Homo sapiens vascular endothelial growth factor D (VEGF-D).
XX
KW vascular endothelial growth factor; VEGF-D; angiogenesis;
KW modification; acceleration; wound healing; tissue; organ;
KW transplants; collateral circulation; infarction; arterial stenosis;
KW coronary artery disease; inhibition; cancer; treatment;
KW diabetic retinopathy; lung disorders; blood circulation;
KW gaseous exchange; chronic obstructive airway disease;
KW intestinal malabsorptive syndrome; biopsy; metastatic risk;
KW detection; diagnosis; congestive heart failure.
XX

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 25, 2002, 16:12:13 : Search time 35.41 Seconds
(without alignments)
46.216 Million cell updates/sec

Title: US-09-854-864-16
Perfect score: 405
Sequence: 1 CPEQYDPLLTGCMCKTI.....DCISCASICGQHPKOCAYFC 67

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 2442594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	393	97.0	166	2	US-08-810-572A-6
2	393	97.0	166	4	US-09-290-333-6
3	393	97.0	233	2	US-08-810-572A-2
4	393	97.0	293	4	US-09-290-333-2
5	76.5	18.9	325	4	US-08-915-795-3
6	76.5	18.9	354	4	US-08-915-795-5
7	74.5	18.4	969	2	US-08-284-941-2
8	74.5	18.4	969	2	US-08-447-642-2
9	74.5	18.4	969	4	US-09-236-503-2
10	74.5	18.4	969	5	PCT-US93-02147A-2
11	70	17.3	1124	1	US-08-323-474-2
12	70	17.3	1124	5	PCT-US93-06093-2
13	69.5	17.2	75	2	US-08-465-380-6
14	69.5	17.2	75	2	US-08-465-380-41
15	69.5	17.2	75	2	US-08-480-478-35
16	69.5	17.2	75	2	US-08-486-397-6
17	69.5	17.2	75	2	US-08-486-397-41
18	69.5	17.2	75	2	US-08-486-399-6
19	69.5	17.2	75	2	US-08-486-399-41
20	69.5	17.2	75	2	US-08-461-965-6
21	69.5	17.2	75	2	US-08-461-965-41
22	69.5	17.2	75	2	US-08-326-110A-35
23	69.5	17.2	75	2	US-08-634-641-6
24	69.5	17.2	75	2	US-08-634-641-41
25	69.5	17.2	75	2	US-09-249-471-6
26	69.5	17.2	75	3	US-09-249-471-41
27	69.5	17.2	75	3	US-09-249-472-6

Sequence 41, Appli
Sequence 6, Appli
Sequence 41, Appli
Sequence 6, Appli
Sequence 41, Appli
Sequence 6, Appli
Sequence 6, Appli
Sequence 6, Appli
Sequence 41, Appli
Sequence 8, Appli
Sequence 37, Appli
Sequence 8, Appli
Sequence 8, Appli
Sequence 8, Appli
Sequence 37, Appli
Sequence 8, Appli
Sequence 8, Appli
Sequence 8, Appli

ALIGNMENTS

RESULT 1
US-08-810-572A-6
; Sequence 6, Application US/08810572A
; Patent No. 5969102

; GENERAL INFORMATION:

; APPLICANT: Bram, Richard J.

; APPLICANT: von Bulow, Gotz

; TITLE OF INVENTION: A LYMPHOCYTE SURFACE RECEPTOR THAT BINDS

; TITLE OF INVENTION: CAML, NUCLEIC ACIDS ENCODING THE SAME AND METHODS OF USE

; TITLE OF INVENTION: THEREOF

; NUMBER OF SEQUENCES: 11

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: David A. Jackson, Esq.

; STREET: 411 Hackensack Ave, Continental Plaza, 4th

; CITY: Hackensack

; STATE: New Jersey

; COUNTRY: USA

; ZIP: 07601

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/810,572A

; FILING DATE: 28-FEB-1997

; CLASSIFICATION: 536

; ATTORNEY/AGENT INFORMATION:

; NAME: Jackson Esq., David A.

; REGISTRATION NUMBER: 26,742

; REFERENCE/DOCKET NUMBER: 1340-1-007

; TELEPHONE: 201-487-5800

; TELEFAX: 201-343-1684

; INFORMATION FOR SEQ ID NO: 6:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 166 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; HYPOTHETICAL: NO

; FRAGMENT TYPE: N-terminal

; ORIGINAL SOURCE:

; ORGANISM: Homo sapiens

; US-08-810-572A-6

Query Match 97.0%; Score 393; DB 2; Length 166;

Best Local Similarity 94.4%; Pred. No. 8e-35; Mismatches 0; Indels 4; Gaps 1;
Matches 67; Conservative 0; Mismatches 0; Indels 4; Gaps 1;
Qy 1 CPEQYWDPLLTGTCMCKTICNHQSQTCAAF-----CRKEQGKFDYDHLRLDRCISCASIC 56
Db 34 CPEQYWDPLLTGTCMCKTICNHQSQTCAAF-----CRKEQGKFDYDHLRLDRCISCASIC 93
Qy 57 GQHPKQCAAYFC 67
Db 94 GQHPKQCAAYFC 104
RESULT 2
US-09-290-333-6
; Sequence 6, Application US/09290333
; Patent No. 6316222
; GENERAL INFORMATION:
; APPLICANT: Bram, Richard J.
; von Bulow, Gotz
; TITLE OF INVENTION: A LYMPHOCYTE SURFACE RECEPTOR THAT BINDS
; CAML, NUCLEIC ACIDS ENCODING THE SAME AND METHODS OF USE
; THEREOF
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th
; Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/290,333
; FILING DATE: 12-Apr-1999
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 1340-1-007 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 166 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-290-333-6

Query Match 97.0%; Score 393; DB 4; Length 166;
Best Local Similarity 94.4%; Pred. No. 8e-35; Mismatches 0; Indels 4; Gaps 1;
Matches 67; Conservative 0; Mismatches 0; Indels 4; Gaps 1;
Qy 1 CPEQYWDPLLTGTCMCKTICNHQSQTCAAF-----CRKEQGKFDYDHLRLDRCISCASIC 56
Db 34 CPEQYWDPLLTGTCMCKTICNHQSQTCAAF-----CRKEQGKFDYDHLRLDRCISCASIC 93
Qy 57 GQHPKQCAAYFC 67
Db 94 GQHPKQCAAYFC 104

RESULT 3
US-08-810-572A-2
; Sequence 2, Application US/08810572A
; Patent No. 5969102
; GENERAL INFORMATION:
; APPLICANT: Bram, Richard J.
; von Bulow, Gotz
; TITLE OF INVENTION: A LYMPHOCYTE SURFACE RECEPTOR THAT BINDS
; CAML, NUCLEIC ACIDS ENCODING THE SAME AND METHODS OF USE
; THEREOF
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/810,572A
; FILING DATE: 28-FEB-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 1340-1-007
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 293 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; US-08-810-572A-2

Query Match 97.0%; Score 393; DB 2; Length 293;
Best Local Similarity 94.4%; Pred. No. 1.4e-34; Mismatches 0; Indels 4; Gaps 1;
Matches 67; Conservative 0; Mismatches 0; Indels 4; Gaps 1;
Qy 1 CPEQYWDPLLTGTCMCKTICNHQSQTCAAF-----CRKEQGKFDYDHLRLDRCISCASIC 56
Db 34 CPEQYWDPLLTGTCMCKTICNHQSQTCAAF-----CRKEQGKFDYDHLRLDRCISCASIC 93
Qy 57 GQHPKQCAAYFC 67
Db 94 GQHPKQCAAYFC 104

RESULT 4
US-09-290-333-2
; Sequence 2, Application US/09290333
; Patent No. 6316222
; GENERAL INFORMATION:
; APPLICANT: Bram, Richard J.
; von Bulow, Gotz
; TITLE OF INVENTION: A LYMPHOCYTE SURFACE RECEPTOR THAT BINDS
; CAML, NUCLEIC ACIDS ENCODING THE SAME AND METHODS OF USE
; THEREOF

```

Query Match          97.0%; Score 393; DB 4; Length 293;
Best Local Similarity 94.4%; Pred. NO. 1.4e-34;
Matches 67; Conservative 0; Mismatches 0; Indels 4; Gaps 1;

QY      1  CPPEQYDPLLGTCMSCKTICNHQSORTCAAFCAFC---CRKEQKGFYDHLRLDCISCASIC 56
      |||||
Db       34  CPPEQYDPLLGTCMSCKTICNHQSORTCAAFCAFCRLSCRKEQKGFYDHLRLDCISCASIC 93
      |||||

QY      57  GQHPKQCAYPEFC 67
      |||||
Db       94  GQHPKQCAYPEFC 104
      |||||

RESULT 5
US-08-915-795-3
; Sequence 3, Application US/08915795
; Patent No. 6235713
; GENERAL INFORMATION:
; APPLICANT: Marc G. ACHEN
; APPLICANT: Andrew F. WILKS
; APPLICANT: Steven A. STACKER
; APPLICANT: Kari ALITALO
; TITLE OF INVENTION: GROWTH FACTOR
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan P.L.L.C.
; STREET: 1200 G Street, NW, Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: United States of America
; ZIP: 20005
; COMPUTER READABLE FORM:

```

TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; TELEX: N/A
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 354 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; TISSUE TYPE: Human Lung
US-08-915-795-5

Query Match 18.9%; Score 76.5; DB 4; Length 354;
Best Local Similarity 24.8%; Pred. No. 0.55;
Matches 28; Conservative 6; Mismatches 22; Indels 57; Gaps 7;

Qy 1 CPEQYWD-----PLIGT-----CCKEKGKFDHLL--RDCISCASICGQHPKQCA 23
Db 222 CPIDMDSNCKVCVQENELAGTDSHLOEPALCGPHMFDRCCECV-CRTPCPKD 280
Qy 24 --SQRTCAAF-----CCCKEKGKFDHLL--RDCISCASICGQHPKQCA 64
Db 281 LIQHPKNCSCFEKSLCTCCQK-----HKLFPDPTCSCEDRCPFHTRPCA 326

RESULT 7
US-284-941-2
; Sequence 2, Application US/08284941
; Patent No. 5863756
; GENERAL INFORMATION:
; APPLICANT: BARR, PHILIP J
; APPLICANT: KIEFER, MICHAEL C
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR PACE 4 AND
; TITLE OF INVENTION: PACE 4.1 GENE AND POLYPEPTIDES IN CELLS
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: COOLEY GODWARD CASTRO HUDDLESON & TATUM
; STREET: FIVE PALO ALTO SQUARE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/284,941
; FILING DATE: 2 August 1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: NEELEY PH.D., RICHARD L.
; REGISTRATION NUMBER: 30092
; REFERENCE/DOCKET NUMBER: CHIR-009/01US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 843-5070
; TELEFAX: (415) 857-0663
; TELEX: 380816 COOLEY PA
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 969 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-284-941-2

Query Match 18.4%; Score 74.5; DB 2; Length 969;
Best Local Similarity 27.8%; Pred. No. 2.3;
Matches 15; Conservative 9; Mismatches 23; Indels 7; Gaps 3;
Qy 14 CMSCKT-ICNHQSORTCAAFCCRCRKEQKGFYDHLRLDCISC---ASICGQHPKQC 63
Db 764 CLSCRRGFYHHQEMNTCVTLG---PAGFYADESQKNCCLKHPCKKCVDEPEKC 814

RESULT 8
US-08-447-642-2
; Sequence 2, Application US/08447642
; Patent No. 5989890
; GENERAL INFORMATION:
; APPLICANT: BARR, PHILIP J
; APPLICANT: KIEFER, MICHAEL C
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR PACE 4 AND
; TITLE OF INVENTION: PACE 4.1 GENE AND POLYPEPTIDES IN CELLS
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: COOLEY GODWARD CASTRO HUDDLESON & TATUM
; STREET: FIVE PALO ALTO SQUARE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/447,642
; FILING DATE: 23-MAY-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/284,941
; FILING DATE: 2 August 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: NEELEY PH.D., RICHARD L.
; REGISTRATION NUMBER: 30092
; REFERENCE/DOCKET NUMBER: CHIR-009/01US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 843-5070
; TELEFAX: (415) 857-0663
; TELEX: 380816 COOLEY PA
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 969 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-447-642-2

Query Match 18.4%; Score 74.5; DB 2; Length 969;
Best Local Similarity 27.8%; Pred. No. 2.3;
Matches 15; Conservative 9; Mismatches 23; Indels 7; Gaps 3;
Qy 14 CMSCKT-ICNHQSORTCAAFCCRCRKEQKGFYDHLRLDCISC---ASICGQHPKQC 63
Db 764 CLSCRRGFYHHQEMNTCVTLG---PAGFYADESQKNCCLKHPCKKCVDEPEKC 814

RESULT 9
US-09-236-503-2
; Sequence 2, Application US/09236503
; Patent No. 6277590
; GENERAL INFORMATION:
; APPLICANT: BARR, PHILIP J
; APPLICANT: KIEFER, MICHAEL C
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR PACE 4 AND 4.1 GENE AND
; TITLE OF INVENTION: Polypeptides in Cells


```
; FILE REFERENCE: CHIR-009/0405
; CURRENT APPLICATION NUMBER: US/09/236,503
; CURRENT FILING DATE: 1999-01-25
; EARLIER APPLICATION NUMBER: 08/447,642
; EARLIER FILING DATE: 1995-05-23
; EARLIER APPLICATION NUMBER: 08/284,941
; EARLIER FILING DATE: 1994-08-02
; EARLIER APPLICATION NUMBER: 07/848,629
; EARLIER FILING DATE: 1992-03-09
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 969
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-236-503-2

Query Match      18.4%; Score 74.5; DB 4; Length 969;
Best Local Similarity 27.8%; Pred. No. 2.3;
Matches 15; Conservative 9; Mismatches 23; Indels 7; Gaps 3;

QY 14 CMCKKT-ICNHQSORTCAAFCCRKEQKGYDHLRDCISC---ASICGQHPKQC 63
Db 764 CLSRRGFYHHQEMTVCVTLCL---PAGFYADESQKNCLKCHPSCKKCVDEPEKC 814

RESULT 10
PCT-US93-02147A-2
; Sequence 2, Application PC/TUS9302147A
; GENERAL INFORMATION:
; APPLICANT: BARR, PHILIP J
; APPLICANT: KIEFER, MICHAEL C
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR PACE 4 AND
; TITLE OF INVENTION: PACE 4.1 GENE AND POLYPEPTIDES IN CELLS
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: COOLEY GOWARD CASTRO HUDDLESON & TATUM
; STREET: FIVE PALO ALTO SQUARE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/02147A
; FILING DATE: 19930309
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/848,629
; FILING DATE: 09-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: NEELEY PH.D., RICHARD L
; REGISTRATION NUMBER: 30092
; REFERENCE/DOCKET NUMBER: CHIR-009/0005
; TELEPHONE: (415) 494-7622
; TELEFAX: (415) 857-0663
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 969 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US93-02147A-2

Query Match      18.4%; Score 74.5; DB 5; Length 969;
```

```
Best Local Similarity 27.8%; Pred. No. 2.3;
Matches 15; Conservative 9; Mismatches 23; Indels 7; Gaps 3;

QY 14 CMCKKT-ICNHQSORTCAAFCCRKEQKGYDHLRDCISC---ASICGQHPKQC 63
Db 764 CLSRRGFYHHQEMTVCVTLCL---PAGFYADESQKNCLKCHPSCKKCVDEPEKC 814

RESULT 11
US-08-323-474-2
; Sequence 2, Application US/08323474
; Patent No. 5447860
; GENERAL INFORMATION:
; APPLICANT: Ziegler, Steven F.
; TITLE OF INVENTION: NOVEL TYROSINE KINASE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/323,474
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/905,600
; FILING DATE: 26-JUN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Seese, Kathryn A.
; REGISTRATION NUMBER: 32,172
; REFERENCE/DOCKET NUMBER: 2609
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 587-0430
; TELEFAX: (206) 233-0644
; TELEX: 756822
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1124 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-323-474-2

Query Match      17.3%; Score 70; DB 1; Length 1124;
Best Local Similarity 26.4%; Pred. No. 8;
Matches 24; Conservative 4; Mismatches 37; Indels 26; Gaps 3;

QY 3 EEQYWDP-----LLGTCL-----SCKTICN-----HQSORTCAAFCCRK 36
Db 212 EAQKNGPECNHCLTACNNGVCHDTGECICPPGFMGRTCCKACELHTFGRTCKERC 271
QY 37 EQGKFYDHLRDCISASICGQHPKQAYFC 67
Db 272 EGCKSYVFCPLDPYGCSCATGWKGLQCNEAC 302

RESULT 12
PCT-US93-06093-2
; Sequence 2, Application PC/TUS9306093
; GENERAL INFORMATION:
; APPLICANT: Ziegler, Steven F.
; TITLE OF INVENTION: NOVEL TYROSINE KINASE
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
```

```

; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/06093
; FILING DATE: 19930625
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/905,600
; FILING DATE: 26-JUN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Seese, Kathryn A.
; REGISTRATION NUMBER: 32,172
; REFERENCE/DOCKET NUMBER: 2609
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 587-0430
; TELEFAX: (206) 233-0644
; TELEX: 756822
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1124 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US93-06093-2

```

Query Match 17.38; Score 70; DB 5; Length 1124;

Best Local Similarity 26.44; Pred. No. 8;

Matches 24; Conservative 4; Mismatches 37; Indels 26; Gaps 3;

```

Qy 3 EQYWDPP-----LLGTCM-----SCKTICN-HQSQRTCAAFCCRK 36
| | | | | | | | | | | | | | | | | | | | | | | | | |
Db 212 EAQKWGPENHLCTACMNGVCHDTGECICPPFGMGFTCEKACELHTFGTKCRCSGQ 271
| | | | | | | | | | | | | | | | | | | | | | | | | |

```

Qy 37 EQGKYDHLRDCISCAISICQGHKQCAAYFC 67

Db 272 BGCKSYVFLPDYPYGCSCATGWKGLQCNEAC 302

RESULT 13

US-08-465-380-6

; Sequence 6, Application US/08465380

; Patent No. 5863894

; GENERAL INFORMATION:

; APPLICANT: George P. Vlasuk, Patric H. Stanssens,

; APPLICANT: Joris H.L. Mensens, Marc J. Lauwereys,

; APPLICANT: Yves R. Laroche, Laurent S. Jespers,

; APPLICANT: Yannick G.J. Gansemans, Matthew Moyle,

; APPLICANT: Peter W. Bergum

; TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT

; NUMBER OF SEQUENCES: 356

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Lyon & Lyon

; STREET: 633 West Fifth Street

; CITY: Suite 4700

; STATE: Los Angeles

; COUNTRY: California

; ZIP: 90071

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5" Diskette, 1.44 mb

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: IBM P.C. DOS 5.0

; SOFTWARE: Word Perfect 5.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/465,380

; FILING DATE: June 5, 1995

; CLASSIFICATION: 530

; PRIOR APPLICATION DATA:

; INFORMATION FOR SEQ ID NO: 6:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 75 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; ORIGINAL SOURCE:

; ORGANISM: Ancylostoma caninum

; US-08-465-380-6

; Query Match 17.28; Score 69.5; DB 2; Length 75;

; Best Local Similarity 27.99; Pred. No. 0.69; Indels 11; Gaps 4;

; Matches 19; Conservative 10; Mismatches 28; Indels 11; Gaps 4;

; Qy 2 PE--EQYWDPLLGTGCMSCKTKICNHQSQR--TCAAFCCKRK-----EQGKFDHLLRDCIS 51

; Db 4 PECCGENWLDVCGTKPKCEAKCSEEBEDPICRSFSCPGPAACVCEGDFYRDTVIGDCVK 63

; Qy 52 CASICGQH 59

; Db 64 -EEECDQH 70

; RESULT 14

; US-08-465-380-41

; Sequence 41, Application US/08465380

; Patent No. 5863894

; GENERAL INFORMATION:

; APPLICANT: George P. Vlasuk, Patric H. Stanssens,

; APPLICANT: Joris H.L. Mensens, Marc J. Lauwereys,

; APPLICANT: Yves R. Laroche, Laurent S. Jespers,

; APPLICANT: Yannick G.J. Gansemans, Matthew Moyle,

; APPLICANT: Peter W. Bergum

; TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT

; NUMBER OF SEQUENCES: 356

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Lyon & Lyon

; STREET: 633 West Fifth Street

; CITY: Suite 4700

; STATE: Los Angeles

; COUNTRY: California

; ZIP: 90071

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5" Diskette, 1.44 mb

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: IBM P.C. DOS 5.0

; SOFTWARE: Word Perfect 5.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/465,380

; FILING DATE: June 5, 1995

; CLASSIFICATION: 530

; PRIOR APPLICATION DATA:

; INFORMATION FOR SEQ ID NO: 6:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 75 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; ORIGINAL SOURCE:

; ORGANISM: Ancylostoma caninum

; US-08-465-380-6

RESULT 15
US-08-480-478-35
; Sequence 35, Application US/08480478
; Patent No. 5864009
; GENERAL INFORMATION:
; APPLICANT: GEORGE P. VLASUK; PATRICK ERIC
; APPLICANT: HUGO STANSENS; JORIS HILDA
; APPLICANT: LIEVEN MESENS; MARC JOZEF
; APPLICANT: LAUREEYS; YVES RENE LAROCHE;
; APPLICANT: LAURENT STEPHANE JESPERS; and
; APPLICANT: YANNICK GEORGES JOZEF
; APPLICANT: GANSEMAN
; TITLE OF INVENTION: NEMATODE-EXTRACTED ANTI-
; TITLE OF INVENTION: COAGULANT PROTEIN
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,478
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/326,110
; FILING DATE: 18 OCTOBER 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158

Search completed: June 25, 2002, 16:12:13
Job time: 53 sec

